



Birr, S., Kley, T., & Volgushev, S. (2019). Model assessment for time series dynamics using copula spectral densities: A graphical tool. *Journal of Multivariate Analysis*, 172(209), 122-146.
<https://doi.org/10.1016/j.jmva.2019.03.003>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.jmva.2019.03.003](https://doi.org/10.1016/j.jmva.2019.03.003)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://doi.org/10.1016/j.jmva.2019.03.003> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Model assessment for time series dynamics using copula spectral densities: a graphical tool

Stefan Birr^a, Tobias Kley^b, Stanislav Volgushev^c

^aRuhr Universität Bochum. Research partially supported by project C1 of SFB 823 of the DFG

^bUniversity of Bristol. Research partially supported by the Engineering and Physical Sciences Research Council grant no. EP/L014246/1

^cUniversity of Toronto. Research partially supported by a discovery grant from Natural Sciences and Engineering Research Council of Canada.

Abstract

Finding parametric models that accurately describe the dependence structure of observed data is a central task in the analysis of time series. Classical frequency domain methods provide a popular set of tools for fitting and diagnostics of time series models, but their applicability is seriously impacted by the limitations of covariances as a measure of dependence. Motivated by recent developments of frequency domain methods that are based on copulas instead of covariances, we propose a novel graphical tool that allows to access the quality of time series models for describing dependencies that go beyond linearity. We provide a thorough theoretical justification of our approach and show in simulations that it can successfully distinguish between subtle differences of time series dynamics, including non-linear dynamics which result from GARCH and EGARCH models. We also demonstrate the utility of the proposed tools through an application to modeling returns of the S&P 500 stock market index.

Keywords: Copula, time series, bootstrap, spectral density, frequency domain

1. Introduction

Non-parametric methods provide valuable tools for dependence modeling. If a parametric candidate model is available, we can compare the corresponding estimate with a non-parametric one to evaluate how well the chosen model describes the data. If no candidate model is available, non-parametric techniques can be used to get a first impression of the underlying dependence and inform about potentially suitable parametric models.

In time series analysis, methods that are based on spectral densities and periodograms have a long and successful history. Priestley [1] suggests to use spectral densities as a graphical tool for model validation by comparing the *spectral shape* of a dataset with standard ones from well known parametric models. Tools based on spectral distributions were considered, among others, by Bartlett [2, 3], who proposed to use the normalized cumulative periodogram to assess whether a process is uncorrelated and to detect hidden periodicities. Rigorous tests for the hypothesis $H_0 : f = f_0$ for a fixed f_0 were derived by Anderson [4], while the more general testing problem $H_0 : f \in \mathcal{F}_\theta$, where \mathcal{F}_θ is a parametric class of spectral densities, was considered by Paparoditis [5]. Fan and Zhang [6] consider generalised likelihood ratio tests for the same hypothesis. There also is a rich literature on non-parametric comparison of the (multivariate) spectra of two time series, here some recent references include [7, 8, 9, 10, 11], but this list is by no means complete.

All of the references cited above deal with classical spectral analysis which is based on the autocovariance function and therefore restricted to the aspects of time series dynamics that can be captured by second-order moments. The autocovariance function does provide a complete description of the dependence of Gaussian processes, but it can completely miss dependencies in a non-Gaussian setting. One such example arises in Economics when first order differences of stock market data (more precisely, of log prices) are analyzed. For illustration, Figure 1 shows the autocorrelations of the log-returns X_t and the squared log-returns X_t^2 calculated from the S&P 500 between 2005 and 2010. While the observations X_t appear to be uncorrelated, we can clearly see positive correlation in the squared observations X_t^2 . This shows that X_t in fact exhibits strong dependence, which however can not be described though the autocovariance function and therefore also completely escapes classical spectral analysis.

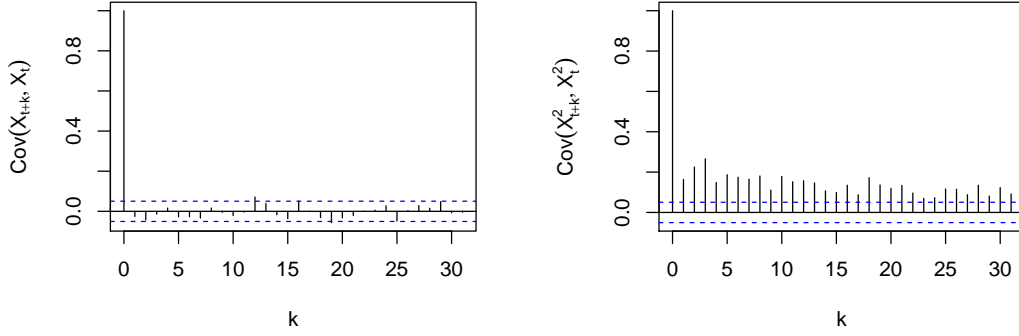


Figure 1: Autocorrelation functions of (daily) log-returns and squared log-returns of the S&P 500 between 2000 and 2005.

This limitation has recently attracted much attention, and new frequency domain tools that can capture non-linear dynamics have been proposed. Pioneering contributions in that direction were made by Hong [12, 13] who introduced *generalized spectra* that are based on joint distributions and joint characteristic functions, respectively. Generalized spectra were later utilized by Hong and Lee [14, 15] and Escanciano [16] to test for the validity of various forms of parametric time series models.

More recently, related approaches were taken by Li [17, 18], who coined the names of *Laplace spectrum* and *Laplace periodogram*. Those ideas were further developed by Hagemann [19] and extended to *cross-spectrum* and spectral *kernel* concepts by Dette et. al [20, 21], who introduced the notion of *copula spectral densities*, Baruník and Kley [22], who introduced *quantile coherency* to measure dependence between economic variables and Birr et. al [23] who consider copula spectra for strictly locally stationary time series.

In the present paper, we utilize copula spectral densities to develop a graphical tool for determining suitable parametric models for time series. We would like to emphasize that our main goal is not to construct yet another test for the hypothesis that a time series is generated by a certain parametric model. Rather, we provide a graphical tool that can indicate whether a chosen model accurately reflects the dependence present in the observed data. By providing useful information about which aspects of the dependence are not captured (if the model is not appropriate) our procedure goes beyond goodness-of-fit tests that merely aim to reject a class of candidate models.

The remaining part of the paper is organized as follows. Section 2.1 contains a summary of basic properties of *copula spectral densities* and provides guidelines for their interpretation. In Section 2.2 we provide details on the proposed algorithm. Section 3 gives a theoretical justification for the graphical approach in the form of a central limit theorem for triangular arrays of processes and Section 4 contains a simulation study and a real data example. All proofs are deferred to the online supplement.

2. Description of Method

2.1. Copula spectral densities: definition and interpretation

We begin by briefly recalling the definition of copula spectral densities. Consider a strictly stationary process $(X_t)_{t \in \mathbb{Z}}$, denote by F its marginal distribution function (which we assume to be continuous), by F_h denote the bivariate distribution functions of (X_{t+h}, X_t) and by C_h the corresponding copulas. Then, the copula spectral density for the process X_t is given by

$$f_{\tau}(\omega) := \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \text{Cov}(\mathbb{I}\{F(X_k) \leq \tau_1\}, \mathbb{I}\{F(X_0) \leq \tau_2\}) e^{-ik\omega} = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} (C_k(\tau_1, \tau_2) - \tau_1 \tau_2) e^{-ik\omega},$$

where $\tau = (\tau_1, \tau_2) \in (0, 1)^2$, $\mathbb{I}\{A\}$ denotes the indicator function of A , and we assume that the terms in the series are absolutely summable. Estimation of copula spectral densities is discussed in the next paragraph, and we begin by

providing more insights about their interpretation. Some of the properties mentioned below have been described in [19, 20, 21], while other parts are new.

We begin by noting that, being based on copulas, the copula spectral density is invariant under strictly increasing marginal transformations of the time series and is thus truly providing information about the temporal dependence structure of the process under consideration. This also implies that the copula spectra of a pair-wise independent time series takes the particularly simple form $f_{(\tau,\eta)}(\omega) \equiv (\tau \wedge \eta - \tau\eta)/2\pi$ which is independent of the marginal distribution.

Next we note that $\omega \mapsto f_\tau(\omega)$ is 2π -periodic for arbitrary $\tau \in [0, 1]^2$ and that f_τ satisfies

$$f_{(\tau_1, \tau_2)}(\omega) = \overline{f_{(\tau_1, \tau_2)}(-\omega)} = \overline{f_{(\tau_2, \tau_1)}(\omega)},$$

where \bar{a} denotes the complex conjugate of a . Those properties imply that the values of $\{f_\tau(\omega) : \tau \in [0, 1]^2, \omega \in [0, \pi]\}$ contain the complete information about the copula spectral density. Even given those restrictions, $f_\tau(\omega)$ is still a complex-valued function of three arguments with each argument taking values in an interval and thus difficult to visualize. One option to get a quick impression of the most important features of the copula spectral density of a given process is to consider all values of $\tau \in \{0.1, 0.5, 0.9\}^2$ and plot the functions $f_\tau(\omega)$ against $\omega \in [0, \pi]$. This requires nine plots which are organized as follows

$$\begin{array}{c|c|c} f_{(0.1, 0.1)}(\omega) & \Im f_{(0.5, 0.1)}(\omega) & \Im f_{(0.9, 0.1)}(\omega) \\ \hline \Re f_{(0.1, 0.5)}(\omega) & f_{(0.5, 0.5)}(\omega) & \Im f_{(0.9, 0.5)}(\omega) \\ \hline \Re f_{(0.1, 0.9)}(\omega) & \Re f_{(0.5, 0.9)}(\omega) & f_{(0.9, 0.9)}(\omega). \end{array}$$

In Figure 2, examples of plots of copula spectral densities corresponding to different parametric models are shown. Those plots will be used to illustrate various properties of copula spectral densities given below.

Begin by observing that $f_{(\tau, \tau)}(\omega)$ is real-valued (for any $\tau \in [0, 1]$ and $\omega \in \mathbb{R}$). It corresponds to the ‘classical’ L^2 spectral density of the clipped process $(I\{F(X_t) \leq \tau\})_{t \in \mathbb{Z}}$ and hence contains information about dynamics of the level-crossing behavior of the process $(X_t)_{t \in \mathbb{Z}}$. A closer look at Figure 2 reveals several interesting features. First, for linear Gaussian processes (i. e., AR(1) and MA(1) with i.i.d. normal innovations) in (a) and (b), the shape of $f_{(\tau, \tau)}(\omega)$ is similar for all values of τ and also similar to the corresponding shape of their L^2 -spectral density. In contrast, the two non-linear models in (c), (d) have copula spectral densities with shape varying across quantile levels. Both models show no dependence at $\tau = 0.5$, which corresponds to an absence of ‘central dependence’. Yet, both processes show a strong dependence (as indicated by sharp peaks for small values of ω) for more extreme quantiles (corresponding to $\tau = 0.1, 0.9$). Note also that the EGARCH model shows an asymmetric dependence with a higher peak at $\tau = 0.1$ compared to $\tau = 0.9$ indicating a stronger serial dependence in the lower tail. In contrast, the dependence in the GARCH model is completely symmetric.

For $\tau \neq \eta$, $f_{(\tau, \eta)}(\omega)$ can be complex-valued. To interpret the real part of $f_{(\tau, \eta)}(\omega)$, note that after a simple computation we obtain for $\tau < \eta$

$$\begin{aligned} \Re f_{(\tau, \eta)}(\omega) &= -\Re \sum_{k \in \mathbb{Z}} e^{-ik\omega} (P(X_k \leq q_\tau, X_0 \geq q_\eta) - \tau(1 - \eta)) \\ &= \tau(1 - \eta) - \sum_{k \geq 1} \cos(k\omega) (P(X_k \leq q_\tau, X_0 \geq q_\eta) - \tau(1 - \eta)) - \sum_{k \geq 1} \cos(k\omega) (P(X_k \geq q_\eta, X_0 \leq q_\tau) - \tau(1 - \eta)). \end{aligned}$$

Hence, the function $\omega \mapsto f_{(\tau, \eta)}(\omega)$ contains information about X_t switching between being below q_τ to above q_η and vice versa. In particular, for τ ‘small’ and η ‘large’ it can be interpreted as describing the dynamics of the process switching between two ‘extreme’ states. Here, the negative peak of $\Re f_{(0.1, 0.9)}$ at small values of ω in (c) indicates that the corresponding GARCH process is likely to switch from a high to a low value (or vice versa), which is exactly what happens in periods of high volatility. Similarly, the positive peak in the same function for (a), (b) corresponds to processes that are unlikely to switch from high to low states immediately, which corresponds to AR(1) or MA(1) dynamics with positive coefficients. It is also interesting to observe that for the two linear processes in (a) and (b) the shapes of $\Re f_{(\tau, \eta)}$ are similar to $f_{(\tau, \tau)}$ for all combinations of τ, η .

The imaginary part of $f_{(\tau, \eta)}(\omega)$ for $\tau < \eta$ takes the form

$$\Im f_{(\tau, \eta)}(\omega) = - \sum_{k \geq 1} \sin(k\omega) (P(X_k \leq q_\tau, X_0 \geq q_\eta) - P(X_k \geq q_\eta, X_0 \leq q_\tau)).$$

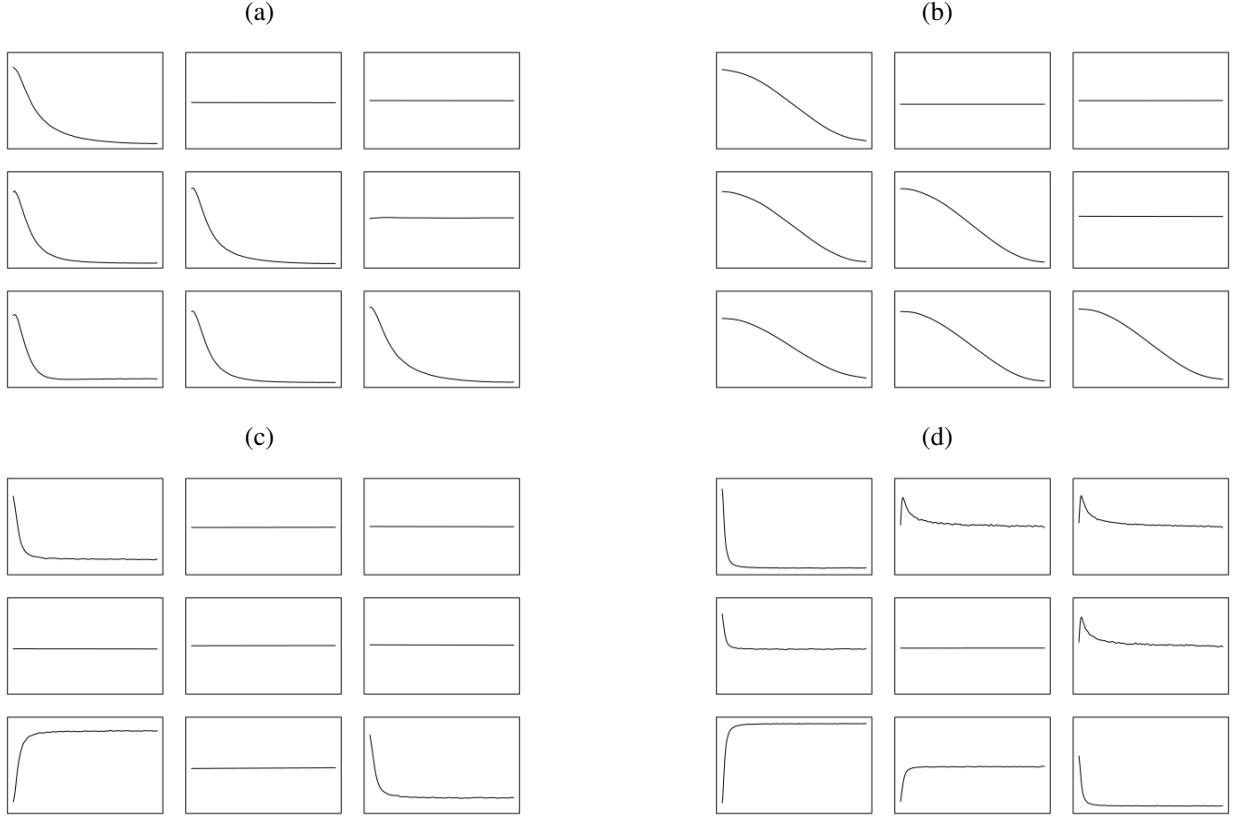


Figure 2: Copula Spectral Densities for $\tau \in \{0.1, 0.5, 0.9\}^2$ of an (a) AR(1), (b) MA(1), (c) GARCH(1,1) and (d) EGARCH(1,1) process.

Note that $\Im f_{(\tau, \eta)} \equiv 0 \Leftrightarrow P(X_k \leq q_\tau, X_0 \geq q_\eta) = P(X_k \geq q_\tau, X_0 \leq q_\eta) \forall k \in \mathbb{Z}$, which shows that $\Im f_{(\tau, \eta)}$ contains information about asymmetry in going from above q_τ to below q_η and vice versa. Non-zero imaginary parts thus indicate time-irreversibility of the dynamics in the observed time series. In particular, if $\Im f_{(\tau, \eta)} \equiv 0$ for all τ, η then this indicates that the process under consideration is pairwise time-reversible, i.e. $C_k(\tau, \eta) = C_{-k}(\tau, \eta)$ for all k, τ, η . The Gaussian linear processes in (a), (b) of Figure 2 are time reversible, which is confirmed by the flat imaginary parts of their copula spectra. It is also noteworthy that the imaginary parts of the processes in (c) and (d) show very different behavior, with clear time-irreversibility for the EGARCH process in (d) and no immediately visible evidence of the same for the GARCH process in (c).

2.2. Graphical tools for model validation

We begin by briefly reviewing estimation of copula spectral densities as discussed in [21] (see also [19] who considered the case $\tau_1 = \tau_2$ and [20] for alternative estimators based on ranks and quantile regression). Given observations X_0, \dots, X_{n-1} we calculate their empirical distribution function $\hat{F}_n(x) := n^{-1} \sum_{t=0}^{n-1} \mathbb{I}\{X_t \leq x\}$ and the copula periodogram

$$I_{\tau, n}(\omega) = \frac{1}{2\pi n} d_{\tau_1, n}(\omega) d_{\tau_2, n}(-\omega),$$

where $\tau = (\tau_1, \tau_2)$ and

$$d_{\tau, n}(\omega) = \sum_{t=0}^{n-1} \mathbb{I}\{\hat{F}_n(X_t) \leq \tau\} e^{-i\omega t}.$$

To obtain a consistent estimator we smooth the copula periodograms over frequencies

$$\hat{f}_\tau(\omega) = \frac{1}{2\pi n} \sum_{s=1}^n W_n(\omega - 2\pi s/n) I_{\tau,n}(2\pi s/n), \quad (1)$$

where W_n denotes a sequence of weighting functions which are specified in assumption (W). Kley et. al [21] proved asymptotic normality (uniformly in τ) of this estimator and computation is possible via the R package **quantspec** (see [24]).

Now, given observations X_1, \dots, X_n we want to decide if that data could have been produced by a parametric model P^{θ_0} where $\{P^\theta : \theta \in \Theta\}$ is a collection of candidate models and $\theta_0 \in \Theta$ is an unknown parameter. To this end we propose to apply parametric bootstrap ideas in the form of Algorithms 1 and 2 given below and on the subsequent page.

Data: Observations X_1, \dots, X_n

Input: Class of parametric models $(P^\theta)_{\theta \in \Theta}$, an estimator $\hat{\theta}$, a collection of frequencies $\omega_1, \dots, \omega_K \in [0, \pi]$, and a quantile level $\tau = (\tau_1, \tau_2)$

Output: Plot comparing copula spectral density estimated from data with ‘typical regions’ created by a parametric bootstrap

```

begin
  Estimate  $\hat{\theta}$  from  $X_1, \dots, X_n$ 
  for  $r$  in  $1:R$  do                                     /* parametric bootstrap */
     $X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$  = simulate from the model  $P^{\hat{\theta}}$ 
     $\hat{f}_\tau^{\hat{\theta},r}(\omega_k)$  = estimated copula spectral density from  $X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$ 
  end
  /* Calculate lower and upper bounds, separately for real and imaginary parts: */
   $l_{\tau,R}^{\Re}(\omega_k) = \alpha/2 - \text{quantile}(\Re \hat{f}_\tau^{\hat{\theta},1}(\omega_k), \dots, \Re \hat{f}_\tau^{\hat{\theta},R}(\omega_k))$ 
   $l_{\tau,R}^{\Im}(\omega_k) = \alpha/2 - \text{quantile}(\Im \hat{f}_\tau^{\hat{\theta},1}(\omega_k), \dots, \Im \hat{f}_\tau^{\hat{\theta},R}(\omega_k))$ 
   $u_{\tau,R}^{\Re}(\omega_k) = (1 - \alpha/2) - \text{quantile}(\Re \hat{f}_\tau^{\hat{\theta},1}(\omega_k), \dots, \Re \hat{f}_\tau^{\hat{\theta},R}(\omega_k))$ 
   $u_{\tau,R}^{\Im}(\omega_k) = (1 - \alpha/2) - \text{quantile}(\Im \hat{f}_\tau^{\hat{\theta},1}(\omega_k), \dots, \Im \hat{f}_\tau^{\hat{\theta},R}(\omega_k))$ 
  /* Estimate the Copula Spectral Density for the data: */
   $\hat{f}_\tau(\omega)$  = estimate the copula spectral density from  $X_1, \dots, X_n$ 
  /* Plot the result */
  plot( $\{\hat{f}_\tau(\omega_k)\}_{k=1,\dots,K}$ )
  plot(Intervals computed from  $(l_{\tau,R}(\omega_k), u_{\tau,R}(\omega_k))_{k=1,\dots,K}$  (separately for real and imaginary parts))
end

```

Algorithm 1: Graphical representation of ‘typical regions’ from a parametric model (with parameter estimated from the data) together with the estimator based on observations.

Algorithm 1 provides a graphical summary of the copula spectral density estimated from data (blue lines) for a few distinct combinations of quantile levels (in the present paper, $(\tau_1, \tau_2) \in \{0.1, 0.5, 0.9\}^2$) as a function of ω together with typical regions (grey areas) that would contain this estimator with probability $1 - \alpha$ if the corresponding class of parametric candidate models was specified correctly (see Proposition 3.1 for a formal statement). One potential concern with Algorithm 1 is that the graphics can become overwhelming if many different quantile levels need to be considered simultaneously. Algorithm 2 below provides a useful supplement to Algorithm 1 which allows to consider many quantile levels at the same time. This necessitates a different graphical representation. The results from Algorithm 2 can be displayed in two different ways. The first provides a summary over all quantile levels $(\tau_1, \tau_2) \in M$ (in the present paper, we choose $M = \{0.05, \dots, 0.95\}^2$) indicating whether the candidate model class produces spectral densities which are compatible with the data for a given frequency but uniformly over quantile levels. If a deviation is detected for a given frequency, a second plot for that particular frequency can be used to determine at which quantile levels the mismatch between the data and the parametric candidate model appears.

Data: Observations X_1, \dots, X_n

Input: Class of parametric models $(P^\theta)_{\theta \in \Theta}$, an estimator $\hat{\theta}$, a frequency $\omega \in [0, \pi]$, and a number K of how many equally spaced quantile levels should be used, quantile level β

Output: Heat-plot of signed p-values indicating whether estimated copula spectral density is within ‘typical regions’ created by a parametric bootstrap

begin

Estimate $\hat{\theta}$ from X_1, \dots, X_n

for r in $1:R$ **do**

/* parametric bootstrap */

$X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$ = simulate from the model $P^{\hat{\theta}}$

$\hat{f}_\tau^{\hat{\theta},r}(\omega)$ = estimated copula spectral density from $X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$, $\forall \tau \in M := \{1/(K+1), \dots, K/(K+1)\}^2$

end

/* Calculate scaling factors, separately for real & imaginary parts: */

Let $l_{\tau,R}^{\Re}(\omega), u_{\tau,R}^{\Re}(\omega)$ denote $\beta/2$ and $1 - \beta/2$ quantile of $\Re \hat{f}_\tau^{\hat{\theta},1}(\omega), \dots, \Re \hat{f}_\tau^{\hat{\theta},R}(\omega)$, respectively (same for \Im)

Define

$$c_{\tau,R}^{\Re}(\omega) = (u_{\tau,R}^{\Re}(\omega) + l_{\tau,R}^{\Re}(\omega))/2, \quad c_{\tau,R}^{\Im}(\omega) = (u_{\tau,R}^{\Im}(\omega) + l_{\tau,R}^{\Im}(\omega))/2$$

and

$$\Delta_{\tau,R}^{\Re}(\omega) = (u_{\tau,R}^{\Re}(\omega) - l_{\tau,R}^{\Re}(\omega))/2, \quad \Delta_{\tau,R}^{\Im}(\omega) = (u_{\tau,R}^{\Im}(\omega) - l_{\tau,R}^{\Im}(\omega))/2 + 10^{-6} I\{u_{\tau,R}^{\Im}(\omega) = l_{\tau,R}^{\Im}(\omega)\}$$

The scaled and centred bootstrap replicate is

$$A_r^{\Re}(\omega) := \max_{\tau=(\tau_1, \tau_2) \in M} |\Re \hat{f}_\tau^{\hat{\theta},R}(\omega) - c_{\tau,R}^{\Re}(\omega)| / \Delta_{\tau,R}^{\Re}(\omega), \quad A_r^{\Im}(\omega) := \max_{\tau=(\tau_1, \tau_2) \in M} |\Im \hat{f}_\tau^{\hat{\theta},R}(\omega) - c_{\tau,R}^{\Im}(\omega)| / \Delta_{\tau,R}^{\Im}(\omega).$$

/* Estimate the Copula Spectral Density for the data: */

$\hat{f}_\tau(\omega)$ = estimate the copula spectral density from X_1, \dots, X_n . Define

$$E_\tau^{\Re}(\omega) := |\Re \hat{f}_\tau(\omega) - c_{\tau,R}^{\Re}(\omega)| / \Delta_{\tau,R}^{\Re}(\omega), \quad E_\tau^{\Im}(\omega) := |\Im \hat{f}_\tau(\omega) - c_{\tau,R}^{\Im}(\omega)| / \Delta_{\tau,R}^{\Im}(\omega)$$

/* Calculate p-values, separately for real and imaginary parts: */

Define \hat{F}_R as the empirical cdf of $\max\{A_1^{\Re}(\omega), A_1^{\Im}(\omega)\}, \dots, \max\{A_R^{\Re}(\omega), A_R^{\Im}(\omega)\}$ and compute

$$p_{\tau,R}^{\Re}(\omega) := 1 - \hat{F}_R(E_\tau^{\Re}(\omega)-), \quad p_{\tau,R}^{\Im}(\omega) := 1 - \hat{F}_R(E_\tau^{\Im}(\omega)-), \quad p_{\min,R}(\omega) := \min_{\tau \in M} \min\{p_{\tau,R}^{\Re}(\omega), p_{\tau,R}^{\Im}(\omega)\}.$$

/* Plot the result */

plot 1: $\omega \mapsto p_{\min,R}(\omega)$; x-axis from $1/R$ to 1, in log-scale. $p_{\min,R}(\omega) = 0$ is indicated by a red circle on the x-axis.)

plot 2: $K \times K$ panels for each ω . The position within each panel corresponds to $\tau \in M$, the symbols used correspond to the magnitude of $p_{\tau,R}^{\Re}(\omega), p_{\tau,R}^{\Im}(\omega)$ (1, 2 and 3 triangles correspond to $p_{\tau,R}^{\Re}(\omega) < 0.05, < 0.01$ and < 0.001 , respectively), and sign of $\Re \hat{f}_\tau(\omega) - c_{\tau,R}^{\Re}(\omega), \Im \hat{f}_\tau(\omega) - c_{\tau,R}^{\Im}(\omega)$ (red triangles facing up indicate a positive and blue triangles facing down indicate a negative value). Information corresponding to $p_{(\tau_i, \tau_j), R}^{\Re}(\omega)$ ($i \geq j$) is in row i column j and information on $p_{(\tau_i, \tau_j), R}^{\Im}(\omega)$ ($i < j$) in row i column j .

end

Algorithm 2: Graphical representation of ‘critical τ s’ from a parametric model (with parameter estimated from the data) together with the estimator based on observations.

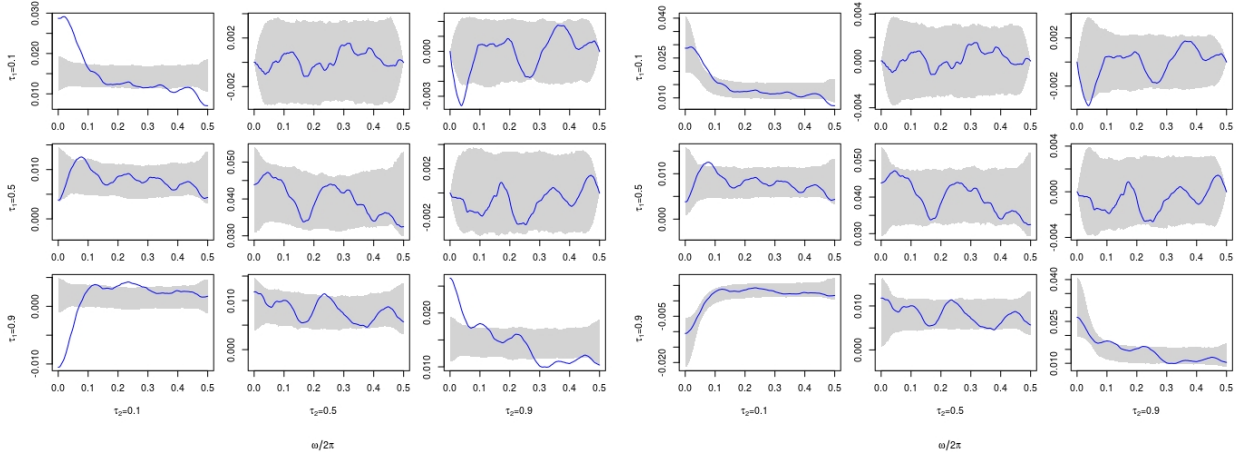


Figure 3: Example using our graphical tool (Algorithm 1) on data generated from a GARCH(1,1) model with $n = 1024$ observations. We are fitting an AR(3) model (left) and a GARCH(1,1) model (right) with $\alpha = 0.05$.

Figure 3 illustrates Algorithm 1 in a simple example. Here the data is a single path simulated from a GARCH(1,1) model, and we considered two classes of parametric models: AR(3) and GARCH(1,1) (the true model is in the latter class, but the parameter was not specified). The blue line shows the estimated copula spectral density \hat{f}_τ (the plot is organized as discussed in Section 2.1; $\Re \hat{f}_\tau$ on/below and $\Im \hat{f}_\tau$ above the diagonal, respectively) and the grey area corresponds to the typical regions for $\alpha = 0.05$ (see Algorithm 1 for details). We clearly see that an AR(3) model is unable to describe the dynamics of a GARCH model, as it fails to capture the dependency in the extreme quantiles ($\tau = (0.1, 0.1), (0.9, 0.9), (0.1, 0.9)$), especially at low frequencies. Considering the true model class (right panel) on the other hand leads to typical regions that almost completely contain the estimated spectrum (note that typical regions are computed pointwise in τ, ω , so the estimator can occasionally be just outside of the boundary of typical regions).

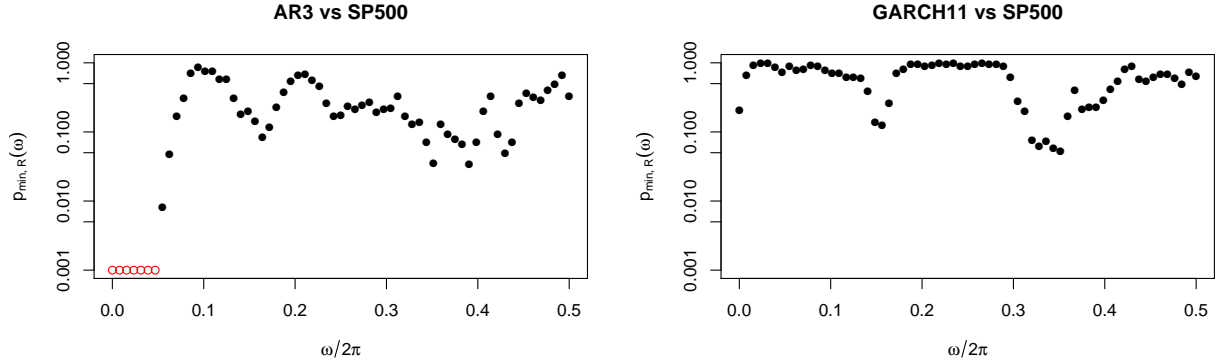


Figure 4: Summary plot produced by Algorithm 2 on data generated from a GARCH(1,1) model with $n = 1024$ observations. We are fitting an AR(3) model (left) and a GARCH(1,1) model (right).

The output of Algorithm 2 for the same data set and models is illustrated in Figures 4 and 5. First, consider the summary plots in Figure 4 with frequencies on the x-axis and $p_{\min,R}(\omega)$ (see Algorithm 2) on the y-axis; for better visibility of very low values the y-axis is in log-scale. By Proposition 3.1, proved below, the values on the y-axis can be interpreted as p-values (uniform in τ and pointwise in ω) of a test for the null hypothesis that the data was generated from a model in the given parametric class against a non-parametric alternative. The left panel of Figure 4 shows the plot corresponding to an AR(3) model class. This plot clearly indicates that the candidate model class does not match the data; this is particularly visible at the lowest frequencies where several p-values in a row are below 0.001. In contrast to that, the right panel which uses the true model class does not show evidence of a misspecified model.

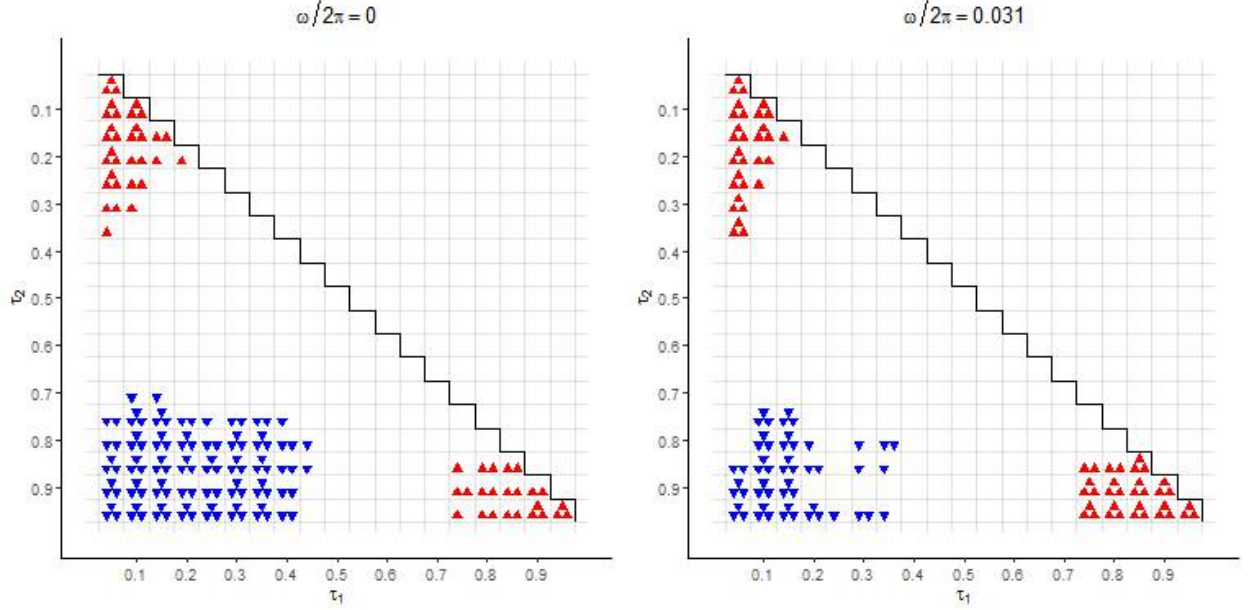


Figure 5: Detailed plots for two frequencies produced by Algorithm 2 on data generated from a GARCH(1,1) model with $n = 1024$ observations with fitted AR(3) model (compare left panel of Figure 4).

The plots in Figure 5 provide more detailed information about the quantile levels at which a mismatch between the data-based spectrum and a spectrum from the candidate parametric model is detected for the frequencies $\omega = 0, \omega = 4\pi/64$. Here, blue triangles facing down indicate that the data-based spectrum is smaller compared to the candidate model spectrum (with 1, 2 and 3 triangles indicating significance at the 5%, 1% and 0.1% level, respectively) while red triangles facing up indicate a data-based spectrum which is larger compared to the candidate model spectrum. The corresponding plots reveal that most of the disagreement between data and model dynamics happens in the real parts of spectra corresponding to quantile levels (τ_1, τ_2) where both τ_1 and τ_2 are either unexpectedly small or unexpectedly large. This confirms the first impression obtained from the 3×3 plots in Figure 3 and provides a more detailed view of the quantiles where data and model spectra disagree.

3. Formal justification of graphical tools

In this section we present a formal justification for the graphical approaches introduced in Section 2.2. Denote by $\Theta \subset \mathbb{R}^d$ a candidate parameter space. For any $\theta \in \Theta$ let $(X_t^\theta)_{t \in \mathbb{Z}}$ be a strictly stationary process distributed according to P^θ . Furthermore, let F^θ denote the cumulative (marginal) distribution function of X_t^θ and denote by F_h^θ the bivariate distribution function of $(X_{t+h}^\theta, X_t^\theta)$. Let C_h^θ denote the copula of $(X_{t+h}^\theta, X_t^\theta)$. We denote the copula spectral density of the process X_t^θ by

$$f_\tau^\theta(\omega) := \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} (C_k^\theta(\tau_1, \tau_2) - \tau_1 \tau_2) e^{-ik\omega}.$$

The corresponding estimator \hat{f}_τ^θ , which is computed from $X_1^\theta, \dots, X_n^\theta$, is denoted by \hat{f}_τ^θ . We make the following technical assumptions

(LC) The copulas C_h^θ are Lipschitz continuous with respect to the parameter θ in a neighborhood of θ_0 uniformly in $\mathcal{T} \subseteq [0, 1]^2$, i.e. there exist constants $\varepsilon > 0, L < \infty$ such that $\|\theta - \theta_0\| \leq \varepsilon$

$$\sup_{h \in \mathbb{Z}} \sup_{\tau \in \mathcal{T}} |C_h^\theta(\tau) - C_h^{\theta_0}(\tau)| \leq L \|\theta - \theta_0\|.$$

(C) For any $p \in \mathbb{N}$ there exist constants $\rho_p \in (0, 1)$ and $K_p < \infty$ such that, for arbitrary intervals $A_1, \dots, A_p \subset \mathbb{R}$ and arbitrary $t_1, \dots, t_p \in \mathbb{Z}$,

$$\sup_{\|\theta - \theta_0\| \leq \varepsilon} |\text{cum}(\mathbb{I}\{X_{t_1}^\theta \in A_1\}, \dots, \mathbb{I}\{X_{t_p}^\theta \in A_p\})| \leq K_p \rho_p^{\max_{i,j} |t_i - t_j|}.$$

(W) The weight function W is real-valued and even with support $[-\pi, \pi]$; moreover it has bounded variation and satisfies $\int W(u) du = 1$. We denote by b_n a sequence of scaling parameters such that $b_n \rightarrow 0$ and $nb_n \rightarrow \infty$, and assume that W_n in (1) takes the form

$$W_n(u) := \sum_{j \in \mathbb{Z}} b_n^{-1} W[b_n^{-1}(u + 2\pi j)].$$

Remark 3.1. Assumption (C) is fulfilled under certain mixing assumptions (see Propositions 3.1 and 3.2 in [21]) and (W) places restriction on the smoothing function which are standard in time series analysis (see for instance page 147 of [25]). (LC) assures that if θ_n is close to θ_0 we also have that the corresponding copula spectral densities are close. Below we show that this assumption is satisfied for ARMA(p,q) processes with normal innovations.

Example 3.1. Let $(X_t^\theta)_{t \in \mathbb{Z}}$ be a strictly stationary ARMA(p,q) process where $\theta = (a_1, \dots, a_p, b_1, \dots, b_q)$ denotes the AR and MA coefficients, that means X_t^θ solves

$$X_t^\theta - \sum_{j=1}^p a_j X_{t-j}^\theta = \epsilon_t + \sum_{i=1}^q b_i \epsilon_{t-i}, \quad (2)$$

where and ϵ_t are centered i.i.d $\mathcal{N}(0, 1)$ random variables. Using the backshift operator B we can write this as $P^\theta(B)X_t^\theta = Q^\theta(B)\epsilon_t$, where P^θ and Q^θ are the polynomials,

$$P^\theta(z) := 1 - a_1 z - \dots - a_p z^p, \quad Q^\theta(z) := 1 + b_1 z + \dots + b_q z^q, \quad z \in \mathbb{C}.$$

To guarantee the existence of a unique strictly stationary and causal solution (see [26]) we assume that Θ is a set such that for all $\theta \in \Theta$ the polynomials P^θ and Q^θ have no common roots and $P^\theta(z)$ only has roots outside the unit circle $\{z \in \mathbb{C} : |z| \leq 1\}$. Under these conditions (LC) holds for any open $\mathcal{T} \subset [0, 1]^2$ and any θ_0 in the interior of Θ . This statement will be proved in Section 5.1

The main result in this section is Proposition 3.1. It implies that, if the parametric model is specified correctly, the intervals $[l_{\tau,R}^\mathfrak{R}(\omega), u_{\tau,R}^\mathfrak{R}(\omega)]$ and $[l_{\tau,R}^\mathfrak{I}(\omega), u_{\tau,R}^\mathfrak{I}(\omega)]$ will (asymptotically) contain the real and imaginary parts of the estimator $\hat{f}_\tau(\omega)$ with given probability α . This provides a formal justification for the graphical approach introduced in Algorithm 1. The second part of Proposition shows that the output of Algorithm 2 can indeed be interpreted as p-values for the null that the class of candidate models contains the true model.

Proposition 3.1. Assume that the data X_1, \dots, X_n are generated from the model P^{θ_0} and let $\hat{\theta}_n$ be a \sqrt{n} -consistent estimator of θ_0 . Let assumptions (LC), (C), (W) hold, assume that $R = R_n \rightarrow \infty$ as $n \rightarrow \infty$ and that there exist constants $k \in \mathbb{N}$ and $\kappa > 0$ with

$$b_n = o(n^{-1/(2k+1)}) \quad b_n n^{1-\kappa} \rightarrow \infty.$$

Then, for $l_{\tau,R_n}(\omega), u_{\tau,R_n}(\omega)$ defined in Algorithm 1 we have, as $n \rightarrow \infty$, for any $\tau \in \mathcal{T}, \omega \in \mathbb{R}$ with $\text{Var}(\Re H_0(\tau, \omega)) \neq 0$,

$$P(l_{\tau,R_n}^\mathfrak{R}(\omega) \leq \Re \hat{f}_\tau(\omega) \leq u_{\tau,R_n}^\mathfrak{R}(\omega)) \rightarrow 1 - \alpha. \quad (3)$$

The same holds for the imaginary parts. If additionally $\min_{\tau \in M} \text{Var}(\Re H_0(\tau, \omega)) > 0$ and $\min_{\tau \in M} \text{Var}(\Im H_0(\tau, \omega)) > 0$ then, for $p_{\tau,R}^\mathfrak{R}(\omega), p_{\tau,R}^\mathfrak{I}(\omega)$ defined in Algorithm 2 we have, as $n \rightarrow \infty$ and any $\omega \in \mathbb{R}$,

$$P\left(\min_{\tau \in M} \min\{p_{\tau,R}^\mathfrak{R}(\omega), p_{\tau,R}^\mathfrak{I}(\omega)\} < \alpha\right) \rightarrow \alpha. \quad (4)$$

The key technical ingredient for proving Proposition 3.1 is given by the following theorem. It provides a generalization of Theorem 3.6 in [21] to a particular kind of triangular array asymptotics. This result is of independent interest, and hence we chose to state it separately.

Theorem 3.1. *Let assumptions (LC), (C), (W) hold, and assume that there exist constants $k \in \mathbb{N}$ and $\kappa > 0$ with*

$$b_n = o(n^{-1/(2k+1)}) \quad b_n n^{1-\kappa} \rightarrow \infty.$$

If $\theta_n = \theta_0 + O(n^{-1/2})$ then, for \mathcal{T} from assumption (LC),

$$\sqrt{nb_n}(\hat{f}_{\tau}^{\theta_n}(\omega) - f_{\tau}^{\theta_0}(\omega) - B_n^{(k)}(\tau, \omega))_{\tau \in \mathcal{T}} \rightsquigarrow H_0(\cdot; \omega)$$

in $\ell^\infty(\mathcal{T})$ where

$$B_n^{(k)}(\tau, \omega) := \begin{cases} \sum_{j=2}^k \frac{b_n^j}{j!} \int_{-\pi}^{\pi} u^j W(u) du (f_{\tau}^{\theta_0})^{(j)}(\omega) & \omega \neq 0 \pmod{2\pi} \\ \frac{n}{2\pi} \tau_1 \tau_2 & \omega = 0 \pmod{2\pi} \end{cases}$$

and $H_0(\cdot; \omega)$ is a complex-valued, centered Gaussian process characterized by

$$\begin{aligned} \text{Cov}(H_0((u_1, v_1); \omega), H_0((u_2, v_2); \omega)) &= 2\pi \int_{-\pi}^{\pi} W^2(u) du \\ &\times \left[(f_{(u_1, u_2)}^{\theta_0}(\omega) f_{(v_1, v_2)}^{\theta_0}(-\omega)) + (f_{(u_1, v_2)}^{\theta_0}(\omega) f_{(v_1, u_2)}^{\theta_0}(-\omega)) \mathbb{I}\{\omega = 0 \pmod{\pi}\} \right]. \end{aligned}$$

4. Simulation study and data example

In this section we present a simulation study and an application to the returns of the S&P 500 stock index between 2000 and 2005 and 1966 and 1970.

4.1. Real data example: S&P 500 returns

In this section we demonstrate how our graphical tools can be utilized to find an appropriate parametric model for a given time series and further provide an example where none of the standard models seem to work. To this end we consider the daily log-returns of the S&P 500 between 03.01.2000 and 30.12.2005 (corresponding to $n = 1508$ observations) as well as between 03.01.1966 and 31.12.1970 (corresponding to $n = 1233$ observations). Throughout this section we use the Epanechnikov kernel for W_n , a moderate bandwidth $b_n = 0.1$ and set $\alpha = 0.05$ in Algorithm 1.

We first consider the daily log-returns of the S&P 500 between 03.01.2000 and 30.12.2005. Assuming for the moment that we have no clue about financial time series we first attempt to fit an AR(3) model. Algorithm 1 with this model class produces Figure 6(a). This clearly shows that an AR(3) manages to capture the “median dependence” but can not account for the strong dependencies observed at $\tau = (0.1, 0.1), (0.9, 0.9)$ and $(0.9, 0.1)$. This is further confirmed in the output produced by Algorithm 2 (see Figure 7(a)). The most basic model which has the potential to model such dependencies is an ARCH(1) model, which is our next candidate. Plot (b) in Figure 6 indeed shows that an ARCH(1) model produces the peaks around frequency $\omega = 0$, but those peaks are not high enough to match the data, this is again confirmed by the summary plot from Algorithm 2 provided in Figure 7(b). Our next try is a GARCH model which was specifically designed to model the types of dependence observed in financial data. Figure 6(c) shows that this model is well suited to reproduce the peaks for $\tau = (0.1, 0.1), (0.9, 0.9)$ and $(0.9, 0.1)$. However, the imaginary parts of the spectra still don’t match the data as can be seen from the part of Figure 6(c) corresponding to $\tau = (0.1, 0.9)$; the mismatch between model and data dynamics is confirmed in the summary plot from Figure 7(c). A closer look at the corresponding detailed plot in the top row of Figure 8 sheds light on the specific combinations of quantile levels for which a significant mismatch occurs. Based on the discussion in Section 2.1 about asymmetric dependence a reasonable model to try is an EGARCH(1,1) model. The output of Algorithm 1 in Figure 6(d) indeed indicates that among all models considered this leads to the best performance, although we still detect slight deviations for some of the imaginary parts. The impression that this model still does not provide a perfect fit is further strengthened by the summary plot in Figure 7(d) where we see a fairly high proportion of p-values below 5% which is much higher than we would expect even after adjusting for multiple testing across frequencies.

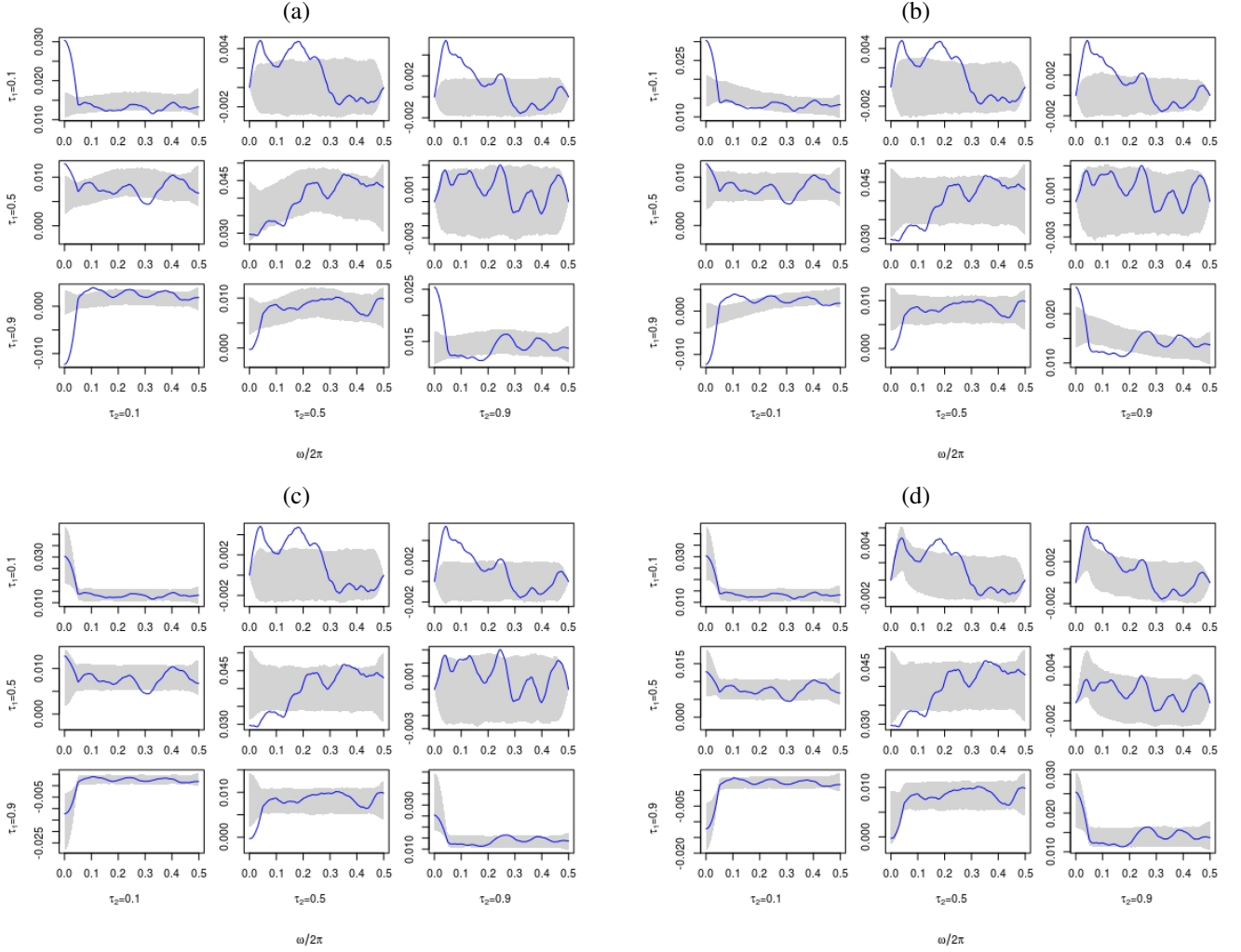


Figure 6: Estimated copula spectral densities based on the daily log-returns of the S&P 500 between 2000 and 2005. Figure displays the plots produced by Algorithm 1 for the following model classes (a) AR(3), (b) ARCH(1), (c) GARCH(1,1) and (d) EGARCH(1,1).

In the final part of this section, we consider the daily log-returns of the S&P 500 between 1966 and 1970. The output of Algorithm 1 for the same four model classes as considered above is depicted in Figure 9 (a)-(d). Interestingly, we find that none of the four model classes provide an adequate description of the dynamics observed in the data since the data contain both - linear type dynamics at the median level, but also strong GARCH-like tail dependencies and EGARCH-like imaginary parts (which are, however, appear to be too steep to be captured by an EGARCH(1,1) model) indicating a strongly asymmetric behaviour of the process going forward and backward in time. The inability of all considered models to capture the dynamics in the data is further confirmed by summary plots from Algorithm 2 as depicted in Figure 10. Additional detailed plots from Algorithm 2 corresponding to specific frequencies are provided in Figure 15 in the online supplement. The middle column corresponding to $\omega = 4 * \pi/64$ confirms that none of the considered models, including the EGARCH model, are able to produce a sufficiently sharp peak in the imaginary part which is observed in the spectrum of the data. The ARCH, GARCH and EGARCH models further struggle to produce the right amount of dependence at central quantile values while the AR(3) process does not have the right kind of dependence in low quantiles.

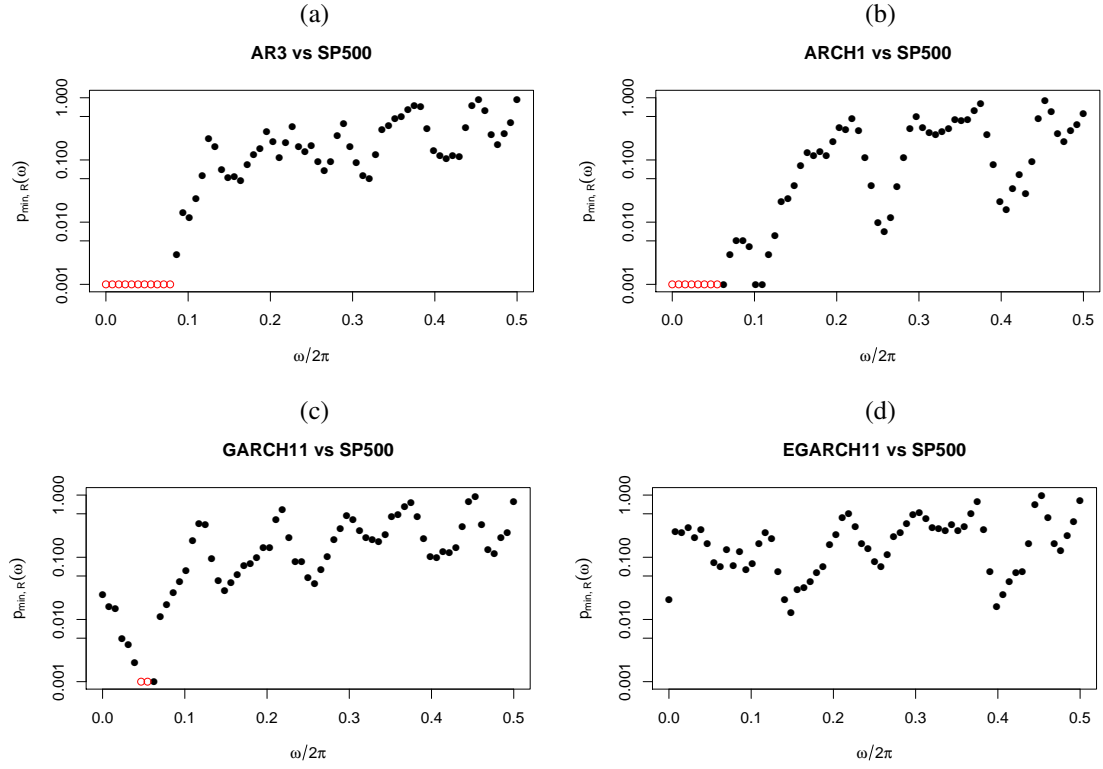


Figure 7: Summary plots produced by Algorithm 2 based on the daily log-returns of the S&P 500 between 2000 and 2005. Figure from top left to bottom right correspond to the following candidate model classes (a) AR(3), (b) ARCH(1), (c) GARCH(1,1) and (d) EGARCH(1,1).

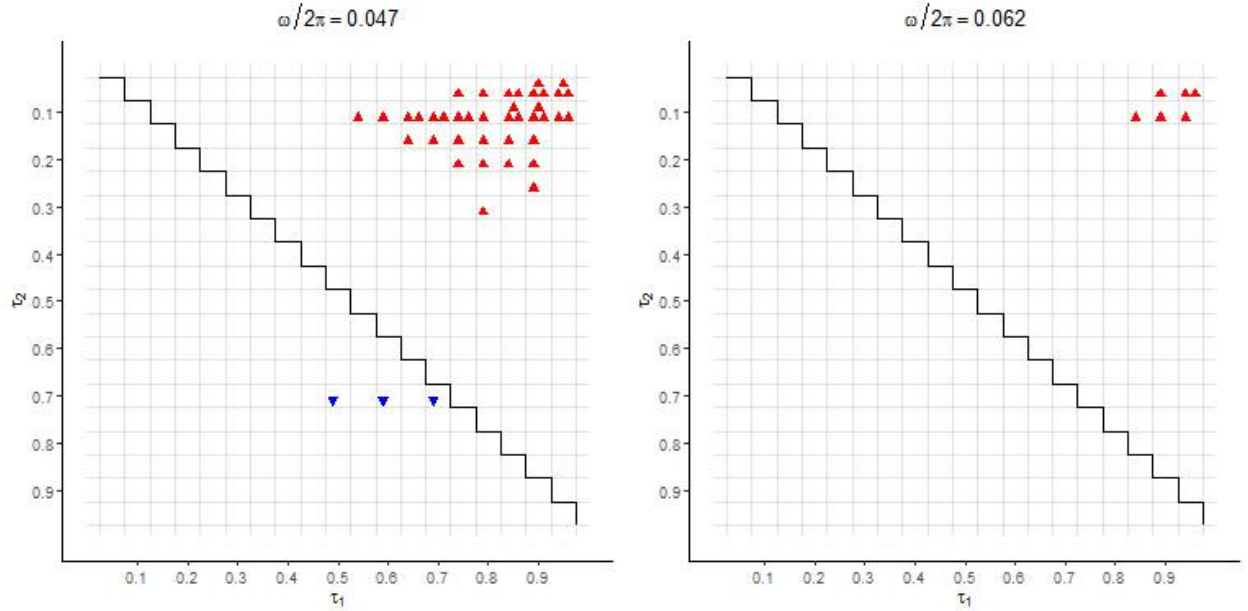


Figure 8: Detailed plots produced by Algorithm 2 at two particular frequencies based on the daily log-returns of the S&P 500 between 2000 and 2005 with GARCH(1,1) as candidate model class.

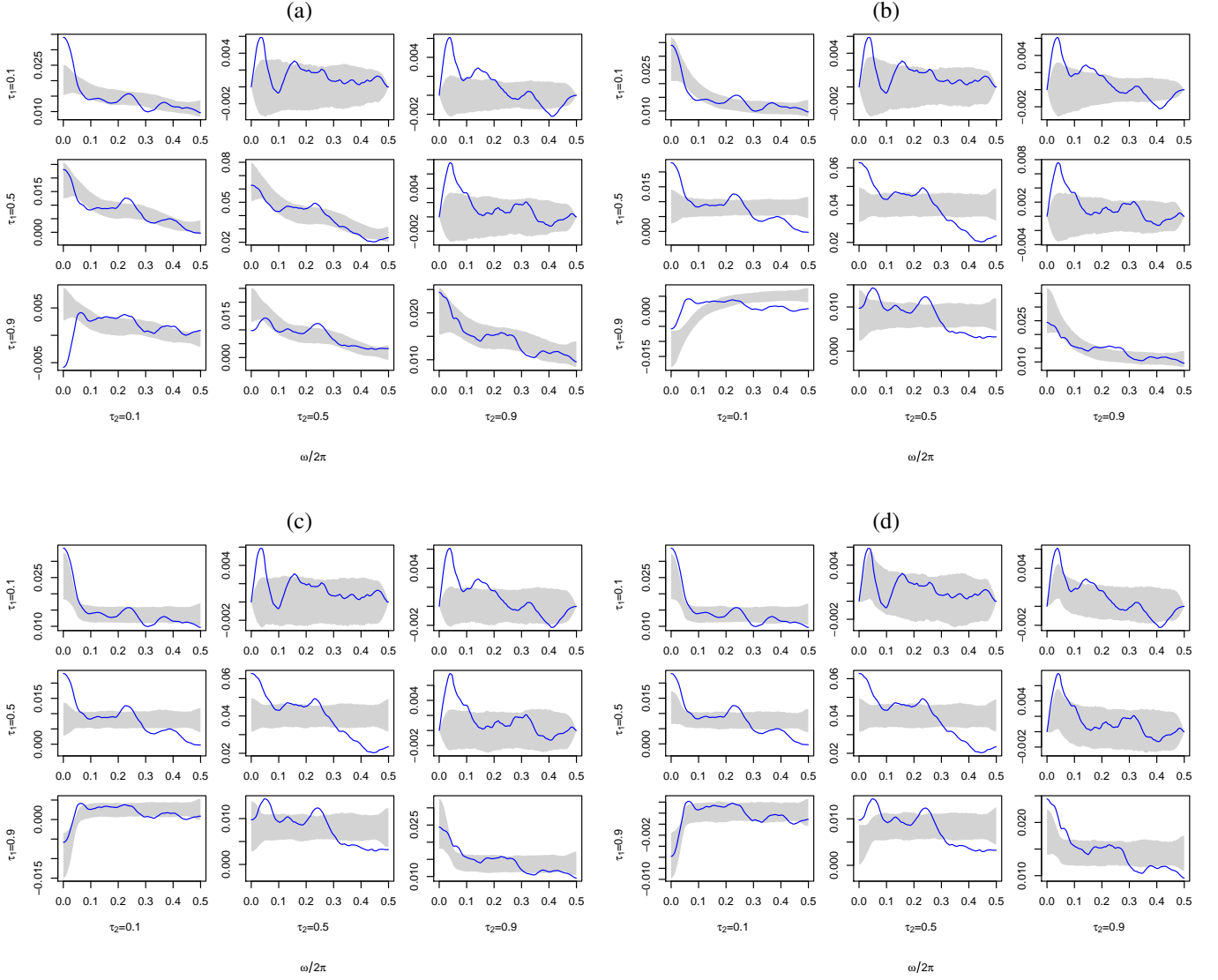


Figure 9: Estimated copula spectral densities based on the daily log-returns of the S&P 500 between 1966 and 1970. Figure displays the plots produced by Algorithm 1 for the following model classes (a) AR(3), (b) ARCH(1), (c) GARCH(1,1) and (d) EGARCH(1,1).

4.2. Simulation study

In this section we illustrate the finite sample properties of Algorithm 1 and Algorithm 2 with simulated data. First, we show that in settings where the data are generated from a model that is contained in the candidate parametric class, the simulated ‘typical regions’ from Algorithm 1 contain the estimated spectral densities with probability $1 - \alpha$ across a range of models, sample sizes and bandwidth parameters (note that this is counted pointwise in τ, ω). To this end, we consider the following three data generating processes.

$$\begin{aligned}
 (a_0) \quad X_t &= 0.1X_{t-1} + 0.8Z_{t-1} + Z_t \\
 (b_0) \quad X_t &= 0.2X_{t-1} - 0.4X_{t-2} + 0.2X_{t-3} + Z_t \\
 (c_0) \quad X_t &= \sigma_t Z_t, \text{ where } \sigma_t^2 = 0.01 + 0.4X_{t-1}^2 + 0.5\sigma_{t-1}^2
 \end{aligned}$$

In each case we simulate time series of length $n = 256, 512, 1024$ and consider the fixed bandwidth parameters $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$. For each possible combination we simulate 1000 repetitions of our algorithm with

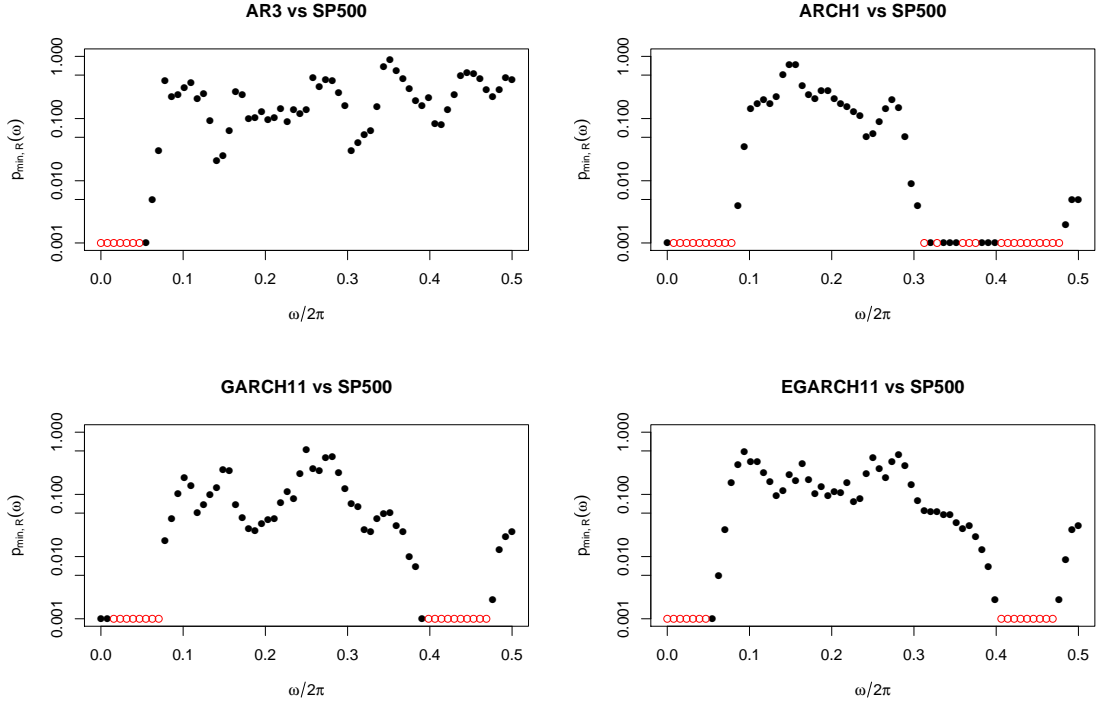


Figure 10: Summary plots produced by Algorithm 2 based on the daily log-returns of the S&P 500 between 1966 and 1970 with GARCH(1,1) as candidate model class.

$\alpha = 0.05$ and the following candidate classes of parametric models (here, θ_j denote unknown parameters of the models)

$$\begin{aligned}
 (P_a) \quad & X_t = \theta_1 X_{t-1} + \theta_2 Z_{t-1} + Z_t, \quad Z_t \sim \mathcal{N}(0, 1) \\
 (P_b) \quad & X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \theta_3 X_{t-3} + Z_t, \quad Z_t \sim \mathcal{N}(0, 1) \\
 (P_c) \quad & X_t = \sigma_t Z_t, \text{ where } \sigma_t^2 = \theta_0 + \theta_1 X_{t-1}^2 + \theta_2 \sigma_{t-1}^2, \quad Z_t \sim \mathcal{N}(0, 1)
 \end{aligned}$$

We use the R packages **QPBoot** [27] which contains useful functions for parametric bootstrap procedures for quantile spectra, **quantspec** [24] to compute the estimators for the copula spectral densities, and **rugarch** [28] to estimate and simulate the GARCH-type models. For each frequency ω we count the number of times the estimated spectral density $\hat{f}_\tau(\omega)$ does not lie in the interval $(l_{\tau,R}(\omega), u_{\tau,R}(\omega))$ (separately for real and imaginary parts). The resulting counts are shown (numbers normalized by 1000) in the left panel of Figures 11, 12 and 13, respectively. We can see that the simulated ‘typical regions’ contain the estimator \hat{f}_τ with prescribed probability across a wide range of scenarios.

Next, we show that the aggregated p-values obtained from Algorithm 2 are calibrated properly. To this end we consider the same models and bandwidth parameters as described above and use 1000 simulation replications to approximate the probabilities

$$P\left(\min_{\tau \in M} \min\{p_{\tau,R}^{\Re}(\omega), p_{\tau,R}^{\Im}(\omega)\} \leq \alpha\right) \quad (5)$$

where we use $M = \{0.05, \dots, 0.95\}^2$ and $R = 1000$. The results corresponding to model $(a_0) - (c_0)$ with candidate model classes $(P_a) - (P_c)$ are depicted in the top three rows of Figure 14 with frequencies on the x-axis and simulated values for the probabilities in (5), with $\alpha = 0.05$, on the y-axis. The plots suggest that the p-values perform as specified in all settings considered.

Next, consider the case when the observations are created by the following models,

- (a₁) $X_t = 0.2X_{t-1} - 0.4X_{t-2} + 0.2X_{t-3} + Z_t$
- (b₁) $X_t = \sigma_t Z_t$, where $\sigma_t^2 = 0.01 + 0.4X_{t-1}^2 + 0.5\sigma_{t-1}^2$
- (c₁) $X_t = \sigma_t Z_t$, where $\ln(\sigma_t^2) = 0.1 + 0.21(|X_{t-1}| - \mathbb{E}|X_{t-1}|) - 0.2X_{t-1} + 0.8\ln(\sigma_{t-1}^2)$,

while the candidate parametric model classes are still $(P_a), (P_b), (P_c)$ and thus are misspecified. The results for Algorithm 1 are depicted in the right panels of Figures 11, 12 and 13, respectively. The plots corresponding to Algorithm 2 are in rows 4–6 of Figure 14.

The results in Figure 11 and row four of Figure 14 show that copula spectral densities are informative for distinguishing different types of linear dynamics (although in this setting any of the classical tests that are tailored to linear models would also be applicable and have excellent power properties). Figure 11 indicates that in this setting copula spectral densities corresponding to $\tau = (0.5, 0.5), (0.1, 0.5), (0.5, 0.9)$ are most informative. This is not surprising since linear dynamics act similarly in all real parts of copula spectral densities and copula spectral densities corresponding to the quantile values mentioned above are easier to estimate (note that for more extreme quantiles only a smaller proportion of the data carry relevant information). Since linear Gaussian processes are time-reversible, the imaginary parts of copula spectra carry no relevant information in this case. Finally, we remark that for this particular data generation process intermediate bandwidth values lead to the most informative results in Figure 11. Row four of Figure 14 additionally shows that aggregating over different frequencies does not lead to a loss in power (despite the uniformity over τ) and in fact improves this probability for the largest bandwidth $b_n = 0.4$.

Figure 12 and row five of Figure 14 show what happens if data are generated by a GARCH model but we attempt to fit their dependence structure by an AR(3) process. In this case the AR(3) model tries to capture the serial correlation of the data, which is zero (so the AR(3) model essentially results in iid data without any serial dependence). This does capture the median dynamics corresponding to $\tau = (0.5, 0.5)$, but completely fails to account for dependence in the more extreme quantiles. This is clearly visible for the real parts of the copula spectral densities corresponding to $\tau = (0.1, 0.1), (0.1, 0.9), (0.9, 0.9)$ with $\tau = (0.1, 0.9)$ leading to the clearest distinction. It is also interesting to observe how different bandwidth values pick up different aspects of the deviation between data and model dynamics. While smaller bandwidth values mainly pick up the sharp peak near zero frequencies, larger bandwidth values also find differences for intermediate frequency values.

The most complicated case that we investigate in our study is to differentiate between a GARCH and an EGARCH process. Results for this are shown in the right panel of Figure 13 and in the bottom row of Figure 14. Both processes have a very similar serial dependence structure as they are uncorrelated but dependent in the extreme quantiles. The difference is that the EGARCH process is asymmetric in the sense, that the dependence is higher in the lower quantiles due to the negative leverage parameter of -0.2 . This difference is subtle and only present in the dependence at large quantiles and hence difficult to pick up and large sample sizes are needed to reliably pick up this distinction. It also turns out that the imaginary part corresponding to $\tau = (0.1, 0.9)$ carries the most information here, with larger bandwidth parameters leading to higher probabilities of detecting relevant differences. The results in the bottom row of Figure 14 additionally show that by aggregating over different quantile levels we are likely to detect deviations between GARCH and an EGARCH processes across a wider range of frequencies. This is due to the fact that for different quantile levels the deviations between the two models are most pronounced at different frequencies.

References

- [1] M. B. Priestley, *Spectral Analysis and Time Series*, Academic Press, 1981.
- [2] M. S. Bartlett, Problèmes de l'analyse spectral des séries temporelles stationnaires, *Publications de l'Institut de statistique de l'Université de Paris III* (3) (1954) 119–134.
- [3] M. S. Bartlett, *An introduction to stochastic processes: with special reference to methods and applications*, CUP Archive, 1978.
- [4] T. W. Anderson, Goodness of fit tests for spectral distributions, *The Annals of Statistics* (1993) 830–847.
- [5] E. Paparoditis, Spectral density based goodness-of-fit tests for time series models, *Scandinavian Journal of Statistics* 27 (1) (2000) 143–176.
- [6] J. Fan, W. Zhang, Generalised likelihood ratio tests for spectral density, *Biometrika* 91 (1) (2004) 195–209. doi:10.1093/biomet/91.1.195.
URL <http://dx.doi.org/10.1093/biomet/91.1.195>
- [7] P. J. Diggle, N. I. Fisher, Nonparametric comparison of cumulative periodograms, *Applied Statistics* (1991) 423–434.
- [8] H. Dette, E. Paparoditis, Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71 (4) (2009) 831–857.
- [9] T. McElroy, S. Holan, A local spectral approach for assessing time series model misspecification, *Journal of Multivariate Analysis* 100 (4) (2009) 604–621. doi:<https://doi.org/10.1016/j.jmva.2008.06.010>.
URL <http://www.sciencedirect.com/science/article/pii/S0047259X08001590>
- [10] C. Jentsch, M. Pauly, Testing equality of spectral densities using randomization techniques, *Bernoulli* 21 (2) (2015) 697–739. doi:10.3150/13-BEJ584.
URL <https://doi.org/10.3150/13-BEJ584>
- [11] J. Chau, H. Ombao, R. von Sachs, Data depth and rank-based tests for covariance and spectral density matrices, Available at arXiv:1706.08289.
- [12] Y. Hong, Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach., *Journal of the American Statistical Association* 94 (1999) 1201–1220.
- [13] Y. Hong, Generalized spectral tests for serial dependence., *Journal of the Royal Statistical Society Ser. B* 62 (2000) 557–574.
- [14] Y. Hong, T.-H. Lee, Diagnostic checking for the adequacy of nonlinear time series models, *Econometric Theory* 19 (6) (2003) 1065–1121. doi:10.1017/S0266466603196089.
- [15] Y. Hong, Y.-J. Lee, Generalized spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form, *The Review of Economic Studies* 72 (2) (2005) 499–541.
- [16] J. C. Escanciano, Goodness-of-fit tests for linear and nonlinear time series models, *Journal of the American Statistical Association* 101 (474) (2006) 531–541. arXiv:<https://doi.org/10.1198/016214505000001050>, doi:10.1198/016214505000001050.
URL <https://doi.org/10.1198/016214505000001050>
- [17] T.-H. Li, Laplace periodogram for time series analysis., *Journal of the American Statistical Association* 103 (2008) 757–768.
- [18] T.-H. Li, Quantile periodograms, *Journal of the American Statistical Association* 107 (2012) 765–776.
- [19] A. Hagemann, Robust spectral analysis, Available at arXiv:1111.1965v2.
- [20] H. Dette, M. Hallin, T. Kley, S. Volgushev, Of copulas, quantiles, ranks and spectra: An L_1 approach to spectral analysis, *Bernoulli* 21 (2015) 781–831.
- [21] T. Kley, S. Volgushev, H. Dette, M. Hallin, Quantile spectral processes: Asymptotic analysis and inference, *Bernoulli* 22 (2016) 1770–1807.
- [22] J. Baruník, T. Kley, Quantile coherency: A general measure for dependence between cyclical economic variables, *The Econometrics Journal* forthcoming.
- [23] S. Birr, S. Volgushev, T. Kley, H. Dette, M. Hallin, Quantile spectral analysis for locally stationary time series, *Journal of the Royal Statistical Society Ser. B*.
- [24] T. Kley, Quantile-Based Spectral Analysis in an Object-Oriented Framework and a Reference Implementation in R: The quantspec Package., *Journal of Statistical Software* 70(3) (2016) 1–27.
- [25] D. R. Brillinger, *Time Series. Data Analysis and Theory*, Holt, Rinehart and Winston, 1975.
- [26] P. J. Brockwell, A. Lindner, Strictly stationary solutions of autoregressive moving average equations, *Biometrika* 97 (2010) 765–772.
- [27] S. Birr, QPboot: Model Validation using Quantile Spectral Analysis and Parametric Bootstrap, R package version 0.2 (2017).
- [28] A. Ghalanos, rugarch: Univariate GARCH models., R package version 1.4-0. (2018).
- [29] A. W. Van der Vaart, *Asymptotic statistics*, Vol. 3, Cambridge university press, 1998.
- [30] A. van der Vaart, J. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer Science & Business Media, 1996.

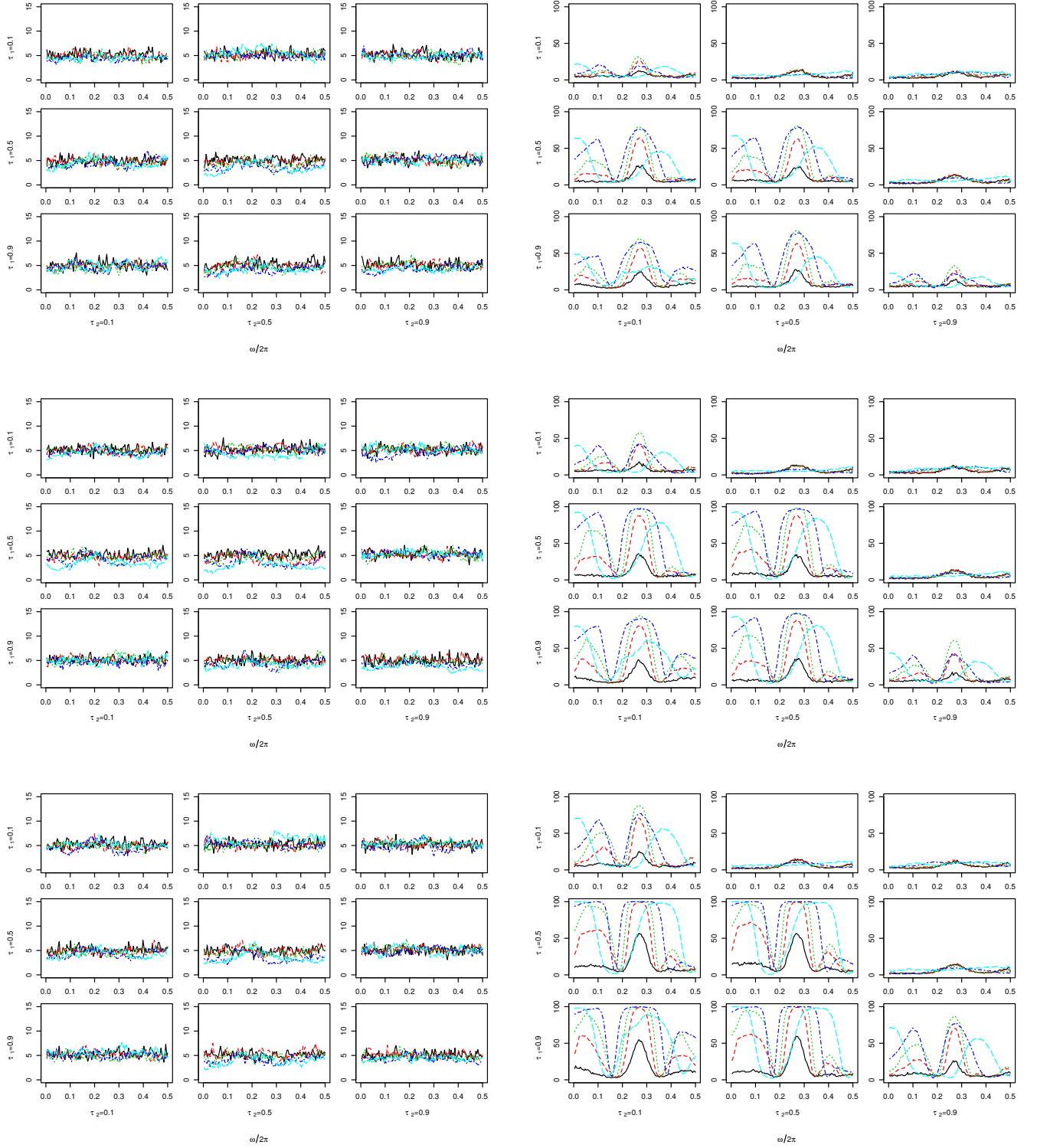


Figure 11: Coverage of the estimator \hat{f} by the critical regions obtained by Algorithm 1. Model class used for the critical regions: P_a (ARMA(1,1)). Data generated according to (a_0) (ARMA(1,1), left panel) or (a_1) (AR(3), right panel). We use $n = 256, 512$ and 1024 observations in the first, second and third row respectively. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively.

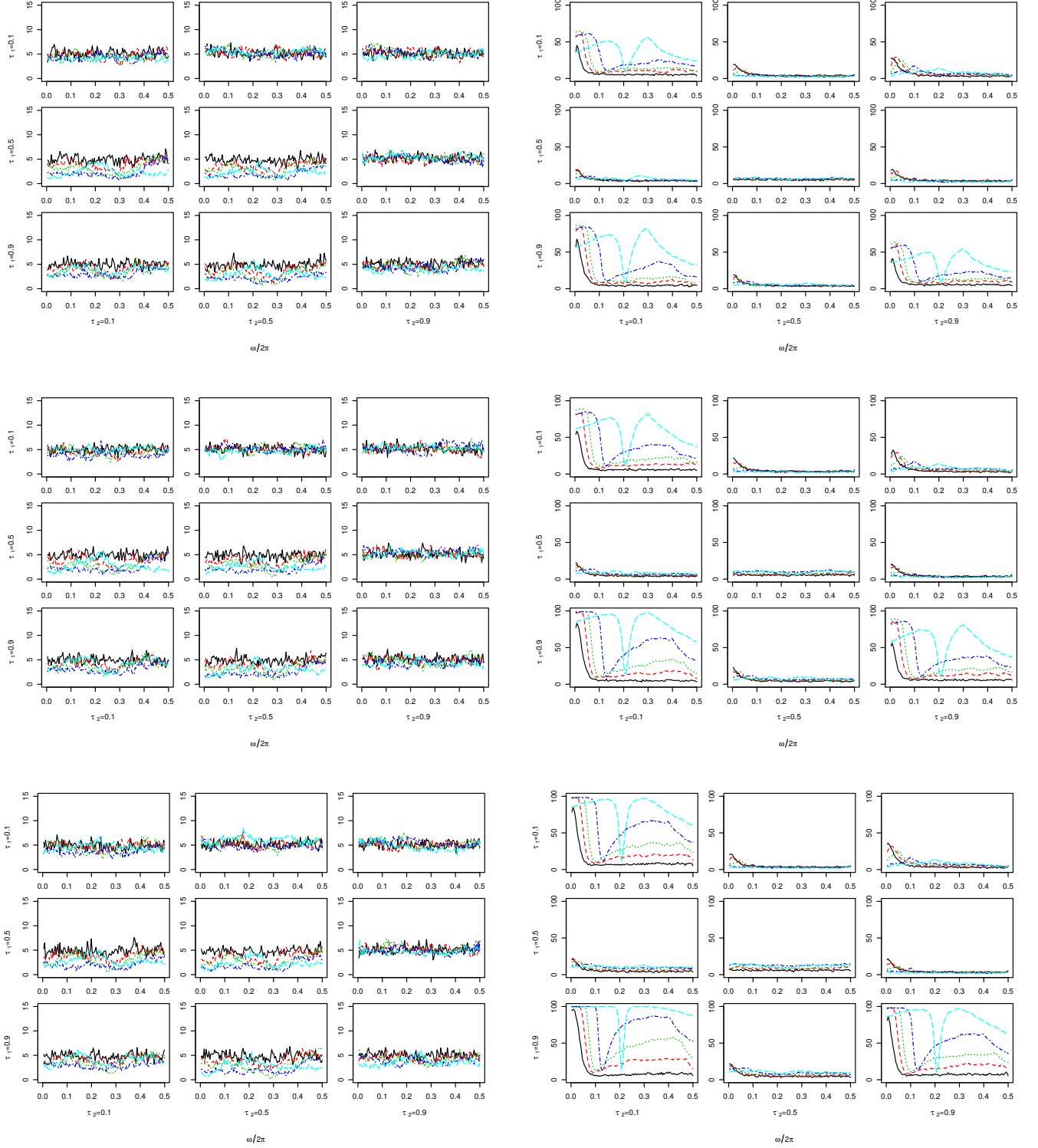


Figure 12: Coverage of the estimator \hat{f} by the critical regions obtained by Algorithm 1. Model class used for the critical regions: P_b (AR(3)). Data generated according to (b_0) (AR(3), left panel) and (b_1) (GARCH(1,1), right panel). We use $n = 256, 512$ and 1024 observations in the first, second and third row respectively. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively.

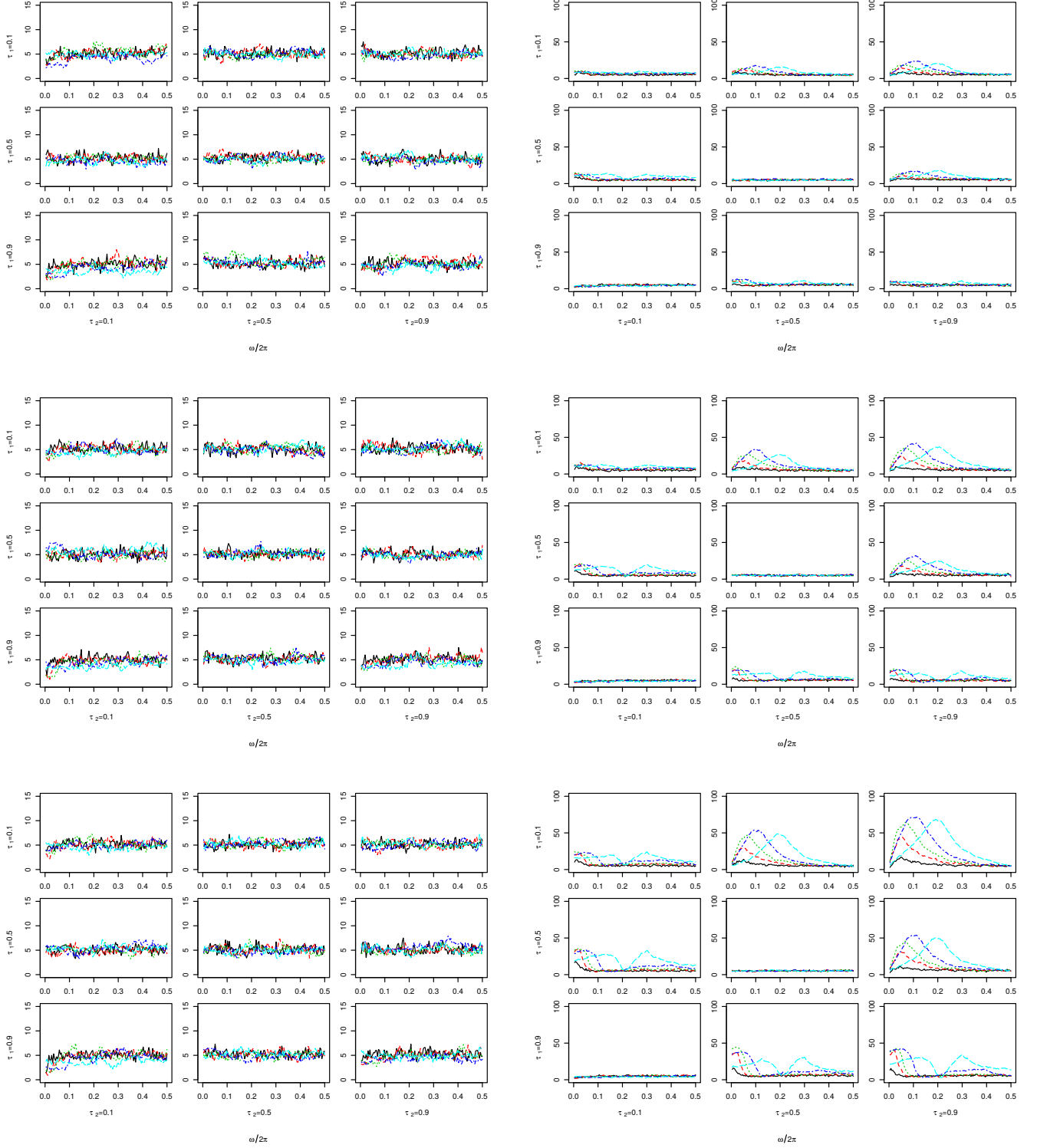
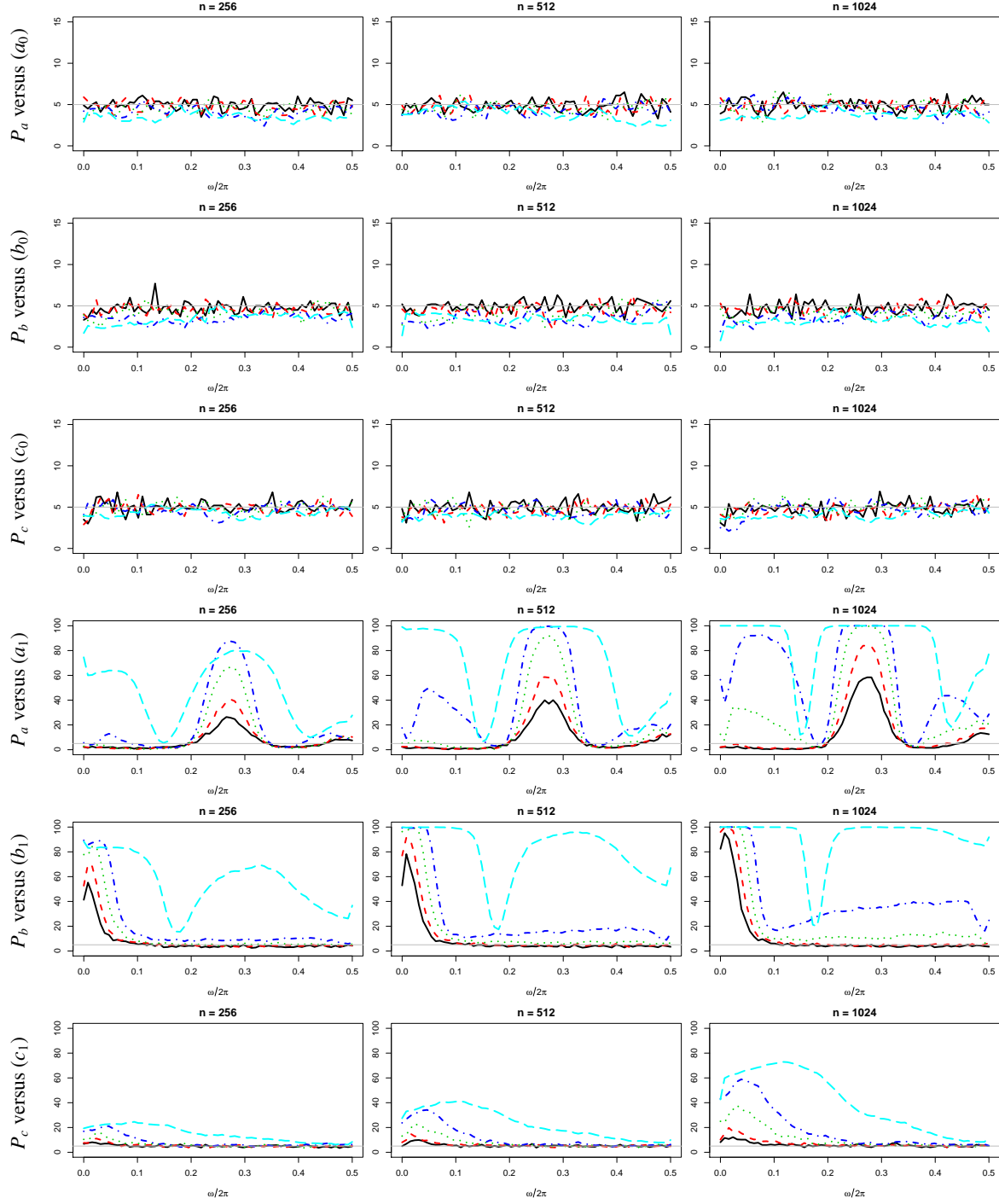


Figure 13: Coverage of the estimator \hat{f} by the critical regions obtained by Algorithm 1. Model class used for the critical regions: P_c (GARCH(1,1)). Data generated according to (c_0) (GARCH(1,1), left panel) and (c_1) (EGARCH(1,1), right panel). We use $n = 256, 512$ and 1024 observations in the first, second and third row respectively. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively.



Online supplement

5. Proofs

5.1. Proof of Example 3.1

Note that under the assumptions made X_t^θ has the representation

$$X_t^\theta = \sum_{k=0}^{\infty} \psi_k^\theta \epsilon_{t-k}$$

where the coefficients are defined by

$$\sum_{k=0}^{\infty} \psi_k^\theta z^k = \frac{Q^\theta(z)}{P^\theta(z)}, \quad |z| \leq 1.$$

By properties of the multivariate normal distribution it suffices to show that for some L

$$\sup_{\mathbf{u} \in \mathbb{R}^2} |F_h^\theta(\mathbf{u}) - F_h^{\theta_0}(\mathbf{u})| \leq \|\theta - \theta_0\|.$$

Applying the triangle inequality we find

$$\sup_{\mathbf{u} \in \mathbb{R}^2} |F_h^\theta(\mathbf{u}) - F_h^{\theta_0}(\mathbf{u})| \leq 2 \sup_{x \in \mathbb{R}} \mathbb{E} [|\mathbb{I}\{X_0^\theta \leq x\} - \mathbb{I}\{X_0^{\theta_0} \leq x\}|]$$

and hence it suffices to show that

$$\sup_{x \in \mathbb{R}} \mathbb{E} [|\mathbb{I}\{X_0^\theta \leq x\} - \mathbb{I}\{X_0^{\theta_0} \leq x\}|] \leq L \|\theta - \theta_0\|.$$

Denote by \mathcal{A}_t the σ -field generated by $\{\epsilon_s | s < t\}$, and by F_ϵ the distribution function of ϵ . This yields

$$\begin{aligned} \sup_{x \in \mathbb{R}} \mathbb{E} (|\mathbb{I}\{X_0^\theta \leq x\} - \mathbb{I}\{X_0^{\theta_0} \leq x\}|) &= \sup_{x \in \mathbb{R}} \mathbb{E} [\mathbb{E} [|\mathbb{I}\{X_t^\theta \leq x\} - \mathbb{I}\{X_t^{\theta_0} \leq x\}| | \mathcal{A}_t]] \\ &= \sup_{x \in \mathbb{R}} \mathbb{E} [\mathbb{E} [|\mathbb{I}\{\epsilon_t \leq x - \sum_{j=1}^{\infty} \psi_j^\theta \epsilon_{t-j}\} - \mathbb{I}\{\epsilon_t \leq x - \sum_{j=1}^{\infty} \psi_j^{\theta_0} \epsilon_{t-j}\}| | \mathcal{A}_t]] \\ &= \sup_{x \in \mathbb{R}} \mathbb{E} [|F_\epsilon(x - \sum_{j=1}^{\infty} \psi_j^\theta \epsilon_{t-j}) - F_\epsilon(x - \sum_{j=1}^{\infty} \psi_j^{\theta_0} \epsilon_{t-j})|] \leq C_1 \mathbb{E} \left| \sum_{j=1}^{\infty} \psi_j^\theta \epsilon_{t-j} - \sum_{j=1}^{\infty} \psi_j^{\theta_0} \epsilon_{t-j} \right| \leq C_2 \sum_{j=1}^{\infty} |\psi_j^\theta - \psi_j^{\theta_0}|. \end{aligned}$$

Finally, we bound the last term above. To shorten notation we write $p_\theta(z) = Q_\theta(z)/P_\theta(z)$. As $P_{\theta_0}(z)$ has no roots on the unit circle, there exist $\eta, \delta > 0$ such that for all $\|\theta - \theta_0\| \leq \eta$

$$P_{\theta_n}(z) \neq 0 \quad \forall z \in \mathbb{C} : |z| < 1 + 2\delta.$$

(Otherwise we could derive a contradiction by using the fact that on \mathbb{C} the locations of roots of a polynomial are a continuous function of the coefficients.) Therefore p_θ is a holomorphic function on $\{z \in \mathbb{C} : |z| \leq 1 + 2\delta\}$ and we can expand $p_\theta(z) = \sum_{j=0}^{\infty} \psi_j^\theta z^j$ and $p_{\theta_0}(z) = \sum_{j=0}^{\infty} \psi_j^{\theta_0} z^j$ with

$$\psi_j^\theta = \frac{1}{2\pi i} \oint_{|\zeta|=1+\delta} \frac{p_\theta(\zeta)}{\zeta^{j+1}} d\zeta, \quad j \in \mathbb{N}_0$$

by Cauchy's differentiation formula. This implies

$$|\psi_j^\theta - \psi_j^{\theta_0}| = \frac{1}{2\pi} \left| \oint_{|\zeta|=1+\delta} \frac{p_\theta(\zeta) - p_{\theta_0}(\zeta)}{\zeta^{j+1}} d\zeta \right|.$$

And with $p_\theta(z) = Q_\theta(z)/P_\theta(z)$ we have that

$$\sup_{|z|=1+\delta} \left| \frac{p_\theta(z) - p_{\theta_0}(z)}{z^{j+1}} \right| \leq \frac{\|\theta - \theta_0\|}{(1+\delta)^{j+1}},$$

which leads to

$$\sum_{j=1}^{\infty} |\psi_j^\theta - \psi_j^{\theta_0}| \leq C_3 \|\theta - \theta_0\| \sum_{j=1}^{\infty} \frac{1}{(1+\delta)^j} =: L \|\theta - \theta_0\|.$$

□

5.2. Proof of Proposition 3.1

We begin by stating a useful Lemma.

Lemma 5.1. *Consider a sequence $r_n = o(1)$ and collection of distribution functions $F_{n,\xi}$ indexed by $\xi \in \Xi, n \in \mathbb{N}$ such that for any deterministic sequence ξ_n in Ξ with $\xi_n = \xi_0 + O(r_n)$ we have $F_{n,\xi_n} \rightsquigarrow F$ for some distribution function F . Then, for any sequence of random variables $\hat{\xi}_n$ in Ξ with $\hat{\xi}_n = \xi_0 + O_P(r_n)$ we have: if Y_1, \dots, Y_{m_n} are i.i.d. $F_{n,\hat{\xi}_n}$ conditional on $\hat{\xi}_n$, $m_n \rightarrow \infty$ and q_n denotes the α th sample quantile of Y_1, \dots, Y_{m_n} then $q_n = F^{-1}(\alpha) + o_P(1)$ for any continuity point α of F^{-1} .*

Proof of Lemma 5.1 Let \hat{F}_n denote the empirical cdf of Y_1, \dots, Y_{m_n} . For any fixed $t \in \mathbb{R}$ we have by the conditional Chebycheff inequality

$$P(|\hat{F}_n(t) - F_{n,\hat{\xi}_n}(t)| \geq \varepsilon | \hat{\xi}_n) \leq (4m_n \varepsilon^2)^{-1} \quad a.s.$$

Taking the expectation with respect to $\hat{\xi}_n$ shows that $\hat{F}_n(t) - F_{n,\hat{\xi}_n}(t) = o_P(1)$ since by assumption $m_n \rightarrow \infty$. Next note that for arbitrary $C > 0$

$$P(|F(t) - F_{n,\hat{\xi}_n}(t)| \geq \varepsilon) \leq \mathbb{I} \left\{ \sup_{|\xi - \xi_0| \leq Cr_n} |F(t) - F_{n,\xi}(t)| \geq \varepsilon \right\} + P(|\hat{\xi}_n - \xi_0| \geq Cr_n)$$

We shall first show that the first term on the right-hand side converges to zero (for $n \rightarrow \infty$) for arbitrary $0 < C < \infty$. Suppose this was not true. Then there exists $\delta > 0$, a subsequence $(n_k)_{k \in \mathbb{N}}$, and $\xi_{n_k} \in \Xi$ with $|\xi_{n_k} - \xi_0| \leq Cr_{n_k}$ and $|F(t) - F_{n_k,\xi_{n_k}}(t)| \geq \delta$ for all $k \in \mathbb{N}$. However, by construction $\xi_{n_k} = \xi_0 + O(r_n)$ (for $k \rightarrow \infty$) which contradicts the assumption. Thus for all $C > 0$

$$\limsup_{n \rightarrow \infty} P(|F(t) - F_{n,\hat{\xi}_n}(t)| \geq \varepsilon) \leq \limsup_{n \rightarrow \infty} P(|\hat{\xi}_n - \xi_0| \geq Cr_n).$$

The right-hand side can be made arbitrarily small by choosing C large since $\hat{\xi}_n = \xi_0 + O_P(r_n)$. Thus we have proved $\hat{F}_n(t) = F(t) + o_P(1)$ for all $t \in \mathbb{R}$.

To complete the proof, observe that \hat{F}_n is a sequence of distribution functions and F is a distribution function. Thus a standard argument implies that $\sup_{t \in \mathbb{R}} |\hat{F}_n(t) - F(t)| = o_P(1)$. This implies $\hat{F}_n^{-1}(\alpha) = F^{-1}(\alpha)$ for all α that are continuity points of F^{-1} ; the latter statement follows by the characterization of convergence in probability in term of a.s. convergence along subsequences and Lemma 21.2 in [29]. □

With the preparations above we are ready to prove Proposition 3.1.

We begin with the proof of (3). Recall the setting and notation introduced in the beginning of Section 3. Let $q(\alpha, \theta_0)$ denote the α -quantile of the distribution of $\mathfrak{R}H_0(\tau; \omega)$ (where H_0 denotes the weak limit in Theorem 3.1). Define

$$g_n := \mathfrak{R}(f_\tau^{\theta_0}(\omega) + B_n^{(k)}(\tau, \omega))$$

for $B_n^{(k)}$ from Theorem 3.1 and let

$$Z_n := \sqrt{nb_n}(\mathfrak{R}\hat{f}_\tau(\omega) - g_n).$$

By Theorem 3.1 applied to the sequence $\theta_n \equiv \theta_0$, $Z_n \rightsquigarrow \mathfrak{R}H_0(\tau; \omega)$ with the limit being a centered normal random variable with non-zero variance.

Now consider the setting of Lemma 5.1 with $m_n = R_n, r_n = n^{-1/2}, \hat{\xi}_n = \hat{\theta}, F$ the cdf of $\Re H_0(\tau; \omega)$, $F_{n,\theta}$ the cdf of $\sqrt{nb_n}(\Re \hat{f}_\tau^\theta(\omega) - g_n)$ and $Y_i = \sqrt{nb_n}(\Re \hat{f}_\tau^{\hat{\theta},i}(\omega) - g_n), i = 1, \dots, R_n$. Note that $\hat{\theta}_n = \theta_0 + O_{\mathbb{P}}(n^{-1/2})$ by assumption and $F_{n,\theta_n} \rightsquigarrow F$ for any sequence $\theta_n = \theta_0 + O(n^{-1/2})$ by Theorem 3.1. Hence, all conditions of Lemma 5.1 are satisfied and we obtain

$$\sqrt{nb_n}(\Re l_{\tau,R_n}^\Re(\omega) - g_n) = q(\alpha/2, \theta_0) + o_P(1).$$

Similarly

$$\sqrt{nb_n}(u_{\tau,R_n}^\Re(\omega) - g_n) = q(1 - \alpha/2, \theta_0) + o_P(1).$$

By Slutsky's Lemma

$$Z_n - \sqrt{nb_n}(u_{\tau,R_n}^\Re(\omega) - g_n) \xrightarrow{\mathcal{D}} \Re H_0(\tau; \omega) - q(1 - \alpha/2, \theta_0),$$

and since the distribution of the limit is continuous

$$P(\Re \hat{f}_\tau(\omega) \leq u_{\tau,R_n}^\Re(\omega)) = P(Z_n \leq \sqrt{nb_n}(u_{\tau,R_n}^\Re(\omega) - g_n)) \rightarrow 1 - \alpha/2.$$

Similarly,

$$P(\Re \hat{f}_\tau(\omega) < l_{\tau,R_n}^\Re(\omega)) = P(Z_n < \sqrt{nb_n}(l_{\tau,R_n}^\Re(\omega) - g_n)) \rightarrow \alpha/2.$$

This completes the proof of (3).

Next let us prove (4). Begin by observing that

$$x \mapsto 1 - \hat{F}_R(x-) = \frac{1}{R} \sum_{r=1}^R I\{x \leq \max\{A_r^\Re(\omega), A_r^\Im(\omega)\}\}$$

is non-increasing, so

$$\min_{\tau \in M} \min \{p_{\tau,R}^\Re(\omega), p_{\tau,R}^\Im(\omega)\} = \frac{1}{R} \sum_{r=1}^R I\left\{ \max_{\tau \in M} \max\{E_\tau^\Re(\omega), E_\tau^\Im(\omega)\} \leq \max\{A_r^\Re(\omega), A_r^\Im(\omega)\} \right\}.$$

Define

$$Z_{r,n}(\tau, \omega) := \sqrt{nb_n}(\hat{f}_\tau^{\hat{\theta},r}(\omega) - f_\tau^{\theta_0}(\omega) - B_n^{(k)}(\tau, \omega))$$

where $B_n^{(k)}$ is defined in Theorem 3.1. Define

$$\begin{aligned} \tilde{l}_{\tau,R}^\Re(\omega) &= \beta/2 - \text{quantile}(\Re Z_{r,n}(\tau, \omega), \dots, \Re Z_{r,n}(\tau, \omega)), \\ \tilde{u}_{\tau,R}^\Re(\omega) &= (1 - \beta/2) - \text{quantile}(\Re Z_{r,n}(\tau, \omega), \dots, \Re Z_{r,n}(\tau, \omega)) \end{aligned}$$

and similar for imaginary parts. Let $\tilde{\Delta}_{\tau,R}^\Re, \tilde{c}_{\tau,R}^\Re, \tilde{A}_r^\Re(\omega)$ denote the corresponding versions of $\Delta_{\tau,R}^\Re, c_{\tau,R}^\Re, A_r^\Re(\omega)$ with $\tilde{u}_{\tau,R}^\Re, \tilde{l}_{\tau,R}^\Re, \Re Z_{r,n}(\tau, \omega)$ instead of $u_{\tau,R}^\Re, l_{\tau,R}^\Re, \hat{f}_\tau^{\hat{\theta},r}(\omega)$ and note that by equivariance of quantiles under the given transformations we have

$$\tilde{u}_{\tau,R}^\Re = \sqrt{nb_n}(u_{\tau,R}^\Re - \Re f_\tau^{\theta_0}(\omega) - \Re B_n^{(k)}(\tau, \omega)), \quad \tilde{l}_{\tau,R}^\Re = \sqrt{nb_n}(l_{\tau,R}^\Re - \Re f_\tau^{\theta_0}(\omega) - \Re B_n^{(k)}(\tau, \omega))$$

which implies $\tilde{A}_r^\Re(\omega) \equiv A_r^\Re(\omega)$ after some simple algebra. From Lemma 5.1 we obtain by similar arguments as above (noting that M is finite) that $\tilde{l}_{\tau,R}^\Re(\omega)$ converges to the $\beta/2$ -quantile of the distribution of $\Re H_0(\tau, \omega)$ and $\tilde{u}_{\tau,R}^\Re(\omega)$ converges to the $1 - \beta/2$ -quantile of the distribution of $\Re H_0(\tau, \omega)$ (both convergences are in probability). Since $\Re H_0(\tau, \omega)$ follows a normal distribution with non-zero variance this implies

$$\max_{\tau \in M} |\tilde{\Delta}_{\tau,R}^\Re(\omega) - \Delta_\tau^\Re(\omega)| = o_P(1), \quad \max_{\tau \in M} |\tilde{c}_{\tau,R}^\Re(\omega)| = o_P(1)$$

where

$$\Delta_{\tau}^{\Re}(\omega) := \sigma_{\tau}^{\Re}(\omega) \left\{ \frac{1}{2} - \Phi^{-1}(\beta/2) \right\}$$

and $\sigma_{\tau}^{\Re}(\omega)$ denotes the standard deviation of $\Re H_0(\tau, \omega)$. Similar results hold for the imaginary parts. By a combination of Slutsky's Lemma and the continuous mapping theorem we now obtain from Theorem 3.1 that

$$\max\{\widetilde{A}_1^{\Re}(\omega), \widetilde{A}_1^{\Im}(\omega)\} \xrightarrow{\mathcal{D}} \frac{1}{\frac{1}{2} - \Phi^{-1}(\beta/2)} \max_{\tau \in M} \max \left\{ \frac{\Re H_0(\tau; \omega)}{\sigma_{\tau}^{\Re}(\omega)}, \frac{\Im H_0(\tau; \omega)}{\sigma_{\tau}^{\Im}(\omega)} \right\}. \quad (6)$$

By similar arguments it follows that

$$\max_{\tau \in M} \max\{E_{\tau}^{\Re}(\omega), E_{\tau}^{\Im}(\omega)\} \xrightarrow{\mathcal{D}} \max_{\tau \in M} \max \left\{ \frac{\Re H_0(\tau; \omega)}{\sigma_{\tau}^{\Re}(\omega)}, \frac{\Im H_0(\tau; \omega)}{\sigma_{\tau}^{\Im}(\omega)} \right\} \sim F. \quad (7)$$

Denoting by F_R the cdf of the random variable $\max\{\widetilde{A}_1^{\Re}(\omega), \widetilde{A}_1^{\Im}(\omega)\}$, the uniform Glivenko-Cantelli Theorem (see Theorem 2.8.1 in [30]) implies that $\sup_{x \in \mathbb{R}} |\hat{F}_R(x) - F_R(x)| = o_P(1)$. Together with (6) and continuity of the cdf, say F , of the random variable $\max_{\tau \in M} \max \left\{ \frac{\Re H_0(\tau; \omega)}{\sigma_{\tau}^{\Re}(\omega)}, \frac{\Im H_0(\tau; \omega)}{\sigma_{\tau}^{\Im}(\omega)} \right\}$ (note that the latter is a maximum over a finite number of (dependent) standard normal random variables, hence has a continuous distribution), it follows that $\sup_{x \in \mathbb{R}} |F_R(x) - F(x)| = o(1)$, and hence

$$\frac{1}{R} \sum_{r=1}^R I \left\{ \max_{\tau \in M} \max\{E_{\tau}^{\Re}(\omega), E_{\tau}^{\Im}(\omega)\} \leq \max\{A_r^{\Re}(\omega), A_r^{\Im}(\omega)\} \right\} = 1 - F \left(\max_{\tau \in M} \max\{E_{\tau}^{\Re}(\omega), E_{\tau}^{\Im}(\omega)\} \right) + o_P(1).$$

Now, by (7) and by continuity of F combined with the continuous mapping Theorem and Slutsky's Lemma we finally obtain

$$\min_{\tau \in M} \min \{p_{\tau, R}^{\Re}(\omega), p_{\tau, R}^{\Im}(\omega)\} = \frac{1}{R} \sum_{r=1}^R I \left\{ \max_{\tau \in M} \max\{E_{\tau}^{\Re}(\omega), E_{\tau}^{\Im}(\omega)\} \leq \max\{A_r^{\Re}(\omega), A_r^{\Im}(\omega)\} \right\} \xrightarrow{\mathcal{D}} 1 - U[0, 1].$$

This completes the proof of (4). \square

5.3. Proof of Theorem 3.1

We will make use of the following notation: $U_t^{\theta} := F^{\theta}(X_t^{\theta})$

$$\begin{aligned} d_{\tau, n}^{U, \theta}(\omega) &:= \sum_{t=0}^{n-1} \mathbb{I}\{U_t^{\theta} \leq \tau\} e^{-i\omega t}, \\ I_{(\tau_1, \tau_2), n}^{U, \theta}(\omega) &:= \frac{1}{2\pi n} d_{\tau_1, n}^{U, \theta}(\omega) d_{\tau_2, n}^{U, \theta}(-\omega), \\ \hat{f}_{\tau, n}^{U, \theta}(\omega) &:= \frac{1}{2\pi n} \sum_{s=1}^{n-1} W_n(\omega - 2\pi s/n) I_{\tau, n}^{U, \theta}(2\pi s/n). \end{aligned}$$

Theorem 3.1 follows from the following four statements.

(i) for any fixed $\omega \in \mathbb{R}$ and an arbitrary sequence θ_n in Θ with $\theta_n = \theta_0 + o(1)$ we have for $n \rightarrow \infty$

$$\sqrt{nb_n}(\hat{f}_{\tau, n}^{U, \theta_n}(\omega) - \mathbb{E} \hat{f}_{\tau, n}^{U, \theta_n}(\omega))_{\tau \in \mathcal{T}} \rightsquigarrow H_0(\cdot; \omega) \quad \text{in } \ell^{\infty}(\mathcal{T})$$

(ii) for $n \rightarrow \infty$ we obtain the following result for the expectation

$$\sup_{\substack{\tau \in [0, 1]^2 \\ \omega \in \mathbb{R}}} |\mathbb{E} \hat{f}_{\tau, n}^{U, \theta_0}(\omega) - f_{\tau}^{\theta_0}(\omega) - B_n^{(k)}(\tau, \omega)| = O((nb_n)^{-1}) + o(b_n^k),$$

(iii) For any fixed ω

$$\sup_{\tau \in [0,1]^2} |\hat{f}_{\tau}^{\theta_n}(\omega) - \hat{f}_{\tau,n}^{U,\theta_n}(\omega)| = o_P((nb_n)^{-1/2} + b_n^k),$$

(iv) for any sequence θ_n in Θ with $\theta_n = \theta_0 + O(n^{-1/2})$ we have for $n \rightarrow \infty$

$$\sqrt{nb_n} \sup_{\tau \in \mathcal{T}, \omega \in \mathbb{R}} \left| \mathbb{E} \hat{f}_{\tau,n}^{U,\theta_n}(\omega) - \mathbb{E} \hat{f}_{\tau,n}^{U,\theta_0}(\omega) \right| = o(1).$$

Note that (ii) is proved in Theorem 3.6(ii) in [21] so that it remains to prove (i), (iii), (iv).

5.3.1. Convergence as a process and the proofs of (i)

Throughout this section, let $\Delta_n(\omega) := \sum_{t=0}^{n-1} e^{i\omega t}$ and $\mathcal{F}_n := \{2\pi j/n : j = 1, \dots, n-1\}$. For intervals $A \subset [0, 1]$ and $\omega \in \mathbb{R}$ define

$$d_{A,n}^{U,\theta_n}(\omega) = \sum_{t=0}^{n-1} \mathbb{I}\{U_t^{\theta_n} \in A\} e^{-it\omega}.$$

Let

$$H_n(\tau, \omega) := \sqrt{nb_n} (\hat{f}_{\tau,n}^{U,\theta_n}(\omega) - \mathbb{E} \hat{f}_{\tau,n}^{U,\theta_n}(\omega))$$

and denote by \hat{F}_n^{U,θ_n} the empirical cdf of $U_1^{\theta_n}, \dots, U_n^{\theta_n}$.

We begin by stating several generalizations of results in [21]. The proofs of those results are very similar to the corresponding proofs in [21] and are omitted for the sake of brevity.

The following statement can be proved similarly to Lemma A.4 in [21]: for arbitrary intervals $A_1, \dots, A_p \subset [0, 1]$ define $\varepsilon := \min_{1 \leq j \leq p} \lambda(A_j)$. Then there exist constants C and d that depend only on p and $K_i, \rho_i, i = 2, \dots, p$ from assumption (C) such that for all $\omega_1, \dots, \omega_p \in \mathbb{R}$

$$\left| \text{cum}(d_{A_1,n}^{U,\theta_n}(\omega_1), \dots, d_{A_p,n}^{U,\theta_n}(\omega_p)) \right| \leq C \left(\left| \Delta_n \left(\sum_{i=1}^p \omega_i \right) \right| + 1 \right) \varepsilon (|\log(\varepsilon)| + 1)^d. \quad (8)$$

Utilizing this statement and following the proof of Lemma 1.6 in the online supplement of [21] we find that for any $k \in \mathbb{N}$ there exists a constant d_k such that, as $\delta_n \rightarrow 0$,

$$\sup_{\substack{x, y \in [0,1] \\ |x-y| \leq \delta_n}} \sqrt{n} |\hat{F}_n^{U,\theta_n}(x) - \hat{F}_n^{U,\theta_n}(y) - x + y| = O_{\mathbb{P}}((n^2 \delta_n + n)^{1/2k} (\delta_n |\log \delta_n|^{d_k} + n^{-1})^{1/2}). \quad (9)$$

This equation combined with the arguments in the proof of Lemma A.6 in [21] shows that for any $k \in \mathbb{N}$

$$\sup_{\omega \in \mathcal{F}_n} \sup_{\tau \in [0,1]} |d_{\tau,n}^{U,\theta_n}(\omega)| = O_{\mathbb{P}}(n^{1/2+1/k}). \quad (10)$$

Now (8), (9) and (10) can be used to replace Lemma A.2, Lemma 1.6 (online supplement), and Lemma A.6 from [21] in the proof of Lemma A.7 in [21] to show the following: if $\delta_n = O((nb_n)^{-1/\gamma})$ for some $\gamma \in (0, 1)$ then

$$\sup_{\omega \in \mathbb{R}} \sup_{\substack{\mathbf{u}, \mathbf{v} \in [0,1]^2 \\ \|\mathbf{u}-\mathbf{v}\| \leq \delta_n}} |H_n(\mathbf{u}, \omega) - H_n(\mathbf{v}, \omega)| = o_{\mathbb{P}}(1). \quad (11)$$

We are now ready for the proof of (i). In view of Theorem 1.5.4 and 1.5.7 in [30] it suffices to show

(i1) convergence of the finite-dimensional distributions

$$(H_n(\tau_j, \omega_j))_{j=1, \dots, k} \rightarrow^d (H_0(\tau_j, \omega_j))_{j=1, \dots, k} \quad (12)$$

for any $(\tau_j, \omega_j) \in \mathcal{T} \times \mathbb{R}, j = 1, \dots, k$ and $k \in \mathbb{N}$.

(i2) stochastic equicontinuity: for any $x > 0$ and any $\omega \in \mathbb{R}$

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{\mathbf{u}, \mathbf{v} \in [0,1]^2, \|\mathbf{u} - \mathbf{v}\| < \delta} |H_n(\mathbf{u}, \omega) - H_n(\mathbf{v}, \omega)| > x \right) = 0. \quad (13)$$

Proof of (i2) Apply Lemma A.2 from [21] with $L = 3$ to obtain

$$\sup_{\omega \in \mathbb{R}} \sup_{\|\mathbf{u} - \mathbf{v}\|_1 \leq \epsilon} \mathbb{E} |H_n(\mathbf{u}, \omega) - H_n(\mathbf{v}, \omega)|^6 \leq K \sum_{l=0}^2 \frac{g(\epsilon)^{3-l}}{(nb_n)^l},$$

here condition (A.2) from Lemma A.2 in [21] is satisfied with $g(x) = x(|\log x| + 1)^d$ by (8). With $\Psi(x) := x^6$ the Orlicz norm $\|X\|_\Psi$ coincides with the L_6 norm $\|X\|_6 = (\mathbb{E}|X|^6)^{1/6}$ so that we have, for any $\kappa \in (0, 1)$ and sufficiently small $\|a - b\|_1$,

$$\|H_n(\mathbf{u}, \omega) - H_n(\mathbf{v}, \omega)\|_\Psi \leq C \left(\frac{\|\mathbf{u} - \mathbf{v}\|_1^\kappa}{(nb_n)^2} + \frac{\|\mathbf{u} - \mathbf{v}\|_1^{2\kappa}}{(nb_n)^1} + \|\mathbf{u} - \mathbf{v}\|_1^{3\kappa} \right).$$

To complete the proof of (i2) follow the arguments in the proof of Theorem 3.6, step (i2), in [21]. Replace Lemma A.7 therein by (11) to obtain for all $x, \mu > 0, 2/3 < \gamma < 1$

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{\mathbf{u}, \mathbf{v} \in [0,1]^2} \sup_{\|\mathbf{u} - \mathbf{v}\|_1 < \delta} |H_n(\mathbf{u}, \omega) - H_n(\mathbf{v}, \omega)| > x \right) \leq \left[\frac{8C}{x} \int_0^\mu z^{-2/(3\gamma)} dz \right]^6.$$

(i2) follows since the integral tends to zero for $\mu \rightarrow 0$.

Proof of (i1) we have to show, that for any $\tau_1, \dots, \tau_k \in \mathcal{T}, k \in \mathbb{N}$ where $\tau_i = (\tau_{i1}, \tau_{i2})$ and $\omega_1, \dots, \omega_k \neq 0 \pmod{2\pi}$ all cumulants of $(H_n(\tau_j, \omega_j))_{j=1, \dots, k}$ converge to the corresponding cumulants of $(H_0(\tau_j, \omega_j))_{j=1, \dots, k}$, which by Lemma P4.5 in [25] gives the desired result. By construction

$$\mathbb{E}(H_n(\tau, \omega)) = 0$$

and

$$\text{Cov}(H_n(\tau_1, \omega_1), H_n(\tau_2, \omega_2)) = nb_n \text{Cov}(\hat{f}_{\tau_1, n}^{U, \theta_n}(\omega_1), \hat{f}_{\tau_2, n}^{U, \theta_n}(\omega_2)).$$

Under assumption (C) the random processes $(\mathbb{I}\{U_t^{\theta_n} \leq \tau_{11}\}, \dots, \mathbb{I}\{U_t^{\theta_n} \leq \tau_{k2}\})_{t \in \mathbb{Z}}$ satisfy a uniform version of Assumption 2.6.2(2) in [25] while the weight functions W_n satisfy Assumption 5.6.1 in [25]. A close look at the proof of Theorem 7.4.3 and Corollary 7.4.3 in [25] shows that all proofs go through without change and leads to the representation

$$\begin{aligned} nb_n \text{Cov}(\hat{f}_{\tau_1, n}^{U, \theta_n}(\omega_1), \hat{f}_{\tau_2, n}^{U, \theta_n}(\omega_2)) &= 2\pi \int_{-\pi}^{\pi} W^2(u) du \left[f_{(\tau_{11}, \tau_{21})}^{\theta_n}(\omega_1) f_{(\tau_{12}, \tau_{22})}^{\theta_n}(-\omega_1) \mathbb{I}\{\omega_1 = \omega_2\} \right. \\ &\quad \left. + f_{(\tau_{11}, \tau_{22})}^{\theta_n}(\omega_1) f_{(\tau_{12}, \tau_{21})}^{\theta_n}(-\omega_1) \mathbb{I}\{\omega_1 = 2\pi - \omega_2\} \right] + O(b_n) + O((nb_n)^{-1}). \end{aligned}$$

Next we note that

$$\sup_{\tau \in \mathcal{T}, \omega \in \mathbb{R}} \left| f_{\tau}^{\theta_n}(\omega) - f_{\tau}^{\theta_0}(\omega) \right| \leq \sum_{h \in \mathbb{Z}} \sup_{\tau \in \mathcal{T}} |C_h^{\theta_n}(\tau) - C_h^{\theta_0}(\tau)| = o(1)$$

by dominated convergence. Hence

$$\begin{aligned} nb_n \text{Cov}(\hat{f}_{\tau_1, n}^{U, \theta_n}(\omega_1), \hat{f}_{\tau_2, n}^{U, \theta_n}(\omega_2)) &\rightarrow 2\pi \int_{-\pi}^{\pi} W^2(u) du \left[f_{(\tau_{11}, \tau_{21})}^{\theta_0}(\omega_1) f_{(\tau_{12}, \tau_{22})}^{\theta_0}(-\omega_1) \mathbb{I}\{\omega_1 = \omega_2\} \right. \\ &\quad \left. + f_{(\tau_{11}, \tau_{22})}^{\theta_0}(\omega_1) f_{(\tau_{12}, \tau_{21})}^{\theta_0}(-\omega_1) \mathbb{I}\{\omega_1 = 2\pi - \omega_2\} \right]. \end{aligned}$$

To complete the proof it remains to show that the cumulants of order $K \geq 3$ vanish as n tends to infinity. We have with $\tau_i = (\tau_{i1}, \tau_{i2})$, $1 \leq i \leq K$ that

$$\begin{aligned} \text{cum}(H_n(\tau_1, \omega_1), \dots, H_n(\tau_K, \omega_K)) &= (nb_n)^{K/2} \text{cum}(\hat{f}_{\tau_1, n}^{U, \theta_n}(\omega_1), \dots, \hat{f}_{\tau_K, n}^{U, \theta_n}(\omega_K)) \\ &= (2\pi)^{2K} n^{-3K/2} (b_n)^{K/2} \sum_{s_1=1}^{n-1} \dots \sum_{s_K=1}^{n-1} W_n(\omega_1 - 2\pi s_1/n) \dots W_n(\omega_K - 2\pi s_K/n) \\ &\quad \text{cum}(d_{\tau_{11}, n}^{U, \theta_n}(2\pi s_1/n) d_{\tau_{12}, n}^{U, \theta_n}(-2\pi s_1/n), \dots, d_{\tau_{K1}, n}^{U, \theta_n}(2\pi s_K/n) d_{\tau_{K2}, n}^{U, \theta_n}(-2\pi s_K/n)). \end{aligned}$$

To see that these cumulants tend to zero we will need arguments similar to those used in Step 2 of the proof of Lemma A.2 in [21]. Applying the product Theorem 2.3.2 in [25] to the last cumulant leads to

$$\begin{aligned} \text{cum}(d_{\tau_{11}, n}^{U, \theta_n}(2\pi s_1/n) d_{\tau_{12}, n}^{U, \theta_n}(-2\pi s_1/n), \dots, d_{\tau_{K1}, n}^{U, \theta_n}(2\pi s_K/n) d_{\tau_{K2}, n}^{U, \theta_n}(-2\pi s_K/n)) \\ = \sum_{(\nu_1, \dots, \nu_N)} \prod_{k=1}^N \text{cum}(d_{\tau_{ij}, n}^{U, \theta_n}((-1)^{j+1} s_i 2\pi/n); (i, j) \in \nu_k), \quad (14) \end{aligned}$$

where the sum runs over all indecomposable partitions (ν_1, \dots, ν_N) (see [25] p. 20) of

$$\begin{array}{cc} (1, 1) & (1, 2) \\ \vdots & \vdots \\ (K, 1) & (K, 2). \end{array}$$

Note that an indecomposable partition consists of at most $N \leq K + 1$ sets. Now by (8) the absolute values of those cumulants are bounded by

$$(14) \leq C \sum_{(\nu_1, \dots, \nu_N)} \prod_{k=1}^N \left[\Delta_n \left(\frac{2\pi}{n} \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \right) + 1 \right] = C \sum_{(\nu_1, \dots, \nu_N)} \sum_{I \subset \{1, \dots, N\}} \prod_{k \in I} \Delta_n \left(\frac{2\pi}{n} \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \right)$$

where C is some constant that depends on K and the constants $K_p, \rho_p, p = 1, \dots, 2K$ from assumption (C) only. Furthermore, since

$$\Delta_n \left(\frac{2\pi}{n} \omega \right) = \begin{cases} n & \omega \in n\mathbb{Z} \\ 0 & \omega \notin n\mathbb{Z} \end{cases}$$

we have that for each combination of $\nu = \{\nu_1, \dots, \nu_N\}$ and $I \subset \{1, \dots, N\}$

$$\prod_{k \in I} \Delta_n \left(\frac{2\pi}{n} \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \right) = 0$$

unless

$$\sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \in n\mathbb{Z} \quad \forall \nu_k \in \nu : k \in I.$$

In the latter case

$$\prod_{k \in I} \Delta_n \left(\frac{2\pi}{n} \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \right) = n^{|I|}.$$

Now we can restrict the sum over the indices (s_1, \dots, s_K) to the set

$$S(\nu, I) := \left\{ (s_1, \dots, s_K) \in \{1, \dots, n-1\}^K : \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \in n\mathbb{Z} \quad \forall \nu_k \in \nu : k \in I \right\}.$$

To complete the proof follow the arguments starting at the bottom of page 16 of the online supplementary material in the proof of Lemma A.2 in [21] (note that the supplement states this as *proof of Lemma 7.2*). First, note that $S(\nu, I)$ is

empty for $|I| = K + 1$ and

$$\begin{aligned} \sum_{s_1, \dots, s_K=1}^{n-1} \prod_{m=1}^K |W_n(\omega_m - 2\pi s_m/n)| \prod_{k \in I} \Delta_n \left(\frac{2\pi}{n} \sum_{(i,j) \in \nu_k} (-1)^{j+1} s_i \right) &= \sum_{(s_1, \dots, s_K) \in S(\nu, I)} \prod_{m=1}^K |W_n(\omega_m - 2\pi s_m/n)| n^{|I|} \\ &= O((b_n^{-1})^{|I| - \lfloor |I|/N \rfloor} n^{K - (|I| - \lfloor |I|/N \rfloor)} n^{|I|}), \end{aligned}$$

where the last equality follows from the arguments around equation (1.26) in the online supplement of [21]. Finally, the number of indecomposable partitions N does not depend on n so that $\text{cum}(H_n(\tau_1, \omega_1), \dots, H_n(\tau_K, \omega_K))$ is of order

$$n^{-3K/2} (b_n)^{K/2} \max_{N \leq K} \max_{|I| \leq N} (b_n^{-1})^{|I| - \lfloor |I|/N \rfloor} n^{K - (|I| - \lfloor |I|/N \rfloor)} n^{|I|} = O((nb_n)^{1-K/2}),$$

which tends to zero for $K \geq 3$. \square

5.3.2. Proof of (iii)

Following the proof of Lemma A.3 and the arguments in the end of the proof of Lemma A.4 in [21] and using (C) it is straightforward to prove that $\omega \mapsto f_{\tau}^{\theta}(\omega)$ is infinitely often continuously differentiable (for any $\tau \in (0, 1)^2$ and $\theta \in U_{\varepsilon}(\theta_0)$) and that there exist constants C, d that are independent of τ_1, τ_2, θ with

$$\sup_{\omega \in \mathbb{R}, \theta \in U_{\varepsilon}(\theta_0)} \left| \frac{d^j}{d\omega^j} f_{\tau_1}^{\theta}(\omega) - \frac{d^j}{d\omega^j} f_{\tau_2}^{\theta}(\omega) \right| \leq C \|\tau_1 - \tau_2\|_1 (1 + \log \|\tau_1 - \tau_2\|_1)^d. \quad (15)$$

Moreover, the proof of Lemma A.5 [21] can be modified to obtain (recall the definition of \hat{F}_n^{U, θ_n} in the beginning of Section 5.3.1)

$$\sup_{\tau \in [0, 1]} |(\hat{F}_n^{U, \theta_n})^{-1}(\tau) - \tau| = O_{\mathbb{P}}(n^{-1/2}). \quad (16)$$

As in [21] (A.4) it follows that for any $k \in \mathbb{N}$ we have

$$\sup_{\omega \in \mathbb{R}} \sup_{\tau \in [0, 1]} |d_{\tau, n}(\omega) - d_{(\hat{F}_n^{U, \theta_n})^{-1}(\tau), n}^U| \leq n \sup_{\tau \in [0, 1]} |\hat{F}_n^{U, \theta_n}(\tau) - \hat{F}_n^{U, \theta_n}(\tau-)| \leq O_{\mathbb{P}}(n^{1/(2k)}),$$

where $\hat{F}_n^{U, \theta_n}(\tau_j-) := \lim_{\xi \uparrow 0} \hat{F}_n^{U, \theta_n}(\tau - \xi)$ and the last inequality follows from (9). The remaining part of the proof is analogous to the arguments given in Section A.3 of [21] and details are omitted for the sake of brevity. \square

5.3.3. Proof of (iv)

Begin by observing that for some constant C_W that depends on W only we have for any $\tau \in [0, 1]^2$

$$\begin{aligned} \left| \mathbb{E} \hat{f}_{\tau, n}^{U, \theta_n}(\omega) - \mathbb{E} \hat{f}_{\tau, n}^{U, \theta_0}(\omega) \right| &\leq \frac{1}{2\pi n} \sum_{s=1}^{n-1} \left| W_n(\omega - 2\pi s/n) \left| \mathbb{E}[I_{\tau, n}^{U, \theta_n}(2\pi s/n)] - \mathbb{E}[I_{\tau, n}^{U, \theta_0}(2\pi s/n)] \right| \right| \\ &\leq C_W \max_{\omega \in \mathcal{F}_n} \left| \mathbb{E}[I_{\tau, n}^{U, \theta_n}(\omega)] - \mathbb{E}[I_{\tau, n}^{U, \theta_0}(\omega)] \right| = C_W \max_{\omega \in \mathcal{F}_n} \left| \frac{1}{2\pi n} \sum_{t_1, t_2=0}^{n-1} e^{-i(t_1 - t_2)\omega} (C_{t_2 - t_1}^{\theta_n}(\tau) - C_{t_2 - t_1}^{\theta_0}(\tau)) \right| \\ &\leq \frac{C_W}{2\pi n} \sum_{|t_1| \leq n} \sum_{k \in \mathbb{Z}} |C_k^{\theta_n}(\tau) - C_k^{\theta_0}(\tau)| \leq C_W \sum_{k \in \mathbb{Z}} |C_k^{\theta_n}(\tau) - C_k^{\theta_0}(\tau)|. \end{aligned}$$

Now under (C) we have $|C_k^{\theta_n}(\tau) - C_k^{\theta_0}(\tau)| \leq 2K_2 \rho_2^{|k|}$. Hence, for any fixed $N \in \mathbb{N}$ we have by (LC)

$$\sup_{\tau \in \mathcal{T}} \sum_{k \in \mathbb{Z}} |C_k^{\theta_n}(\tau) - C_k^{\theta_0}(\tau)| \leq \sum_{|k| \leq N} \sup_{\tau \in \mathcal{T}} |C_k^{\theta_n}(\tau) - C_k^{\theta_0}(\tau)| + 4K_2 \sum_{k > N} \rho_2^{|k|} \leq (2N + 1) \|\theta_n - \theta_0\| + \frac{4K_2 \rho_2^{N+2}}{1 - \rho_2}.$$

Now by assumption $\|\theta_n - \theta_0\| = O(n^{-1/2})$, so picking $N = C \log n$ for a constant $C > 0$ such that $4K_2 \rho_2^{C \log n} = o(n^{-1})$ we obtain

$$\sup_{\tau \in \mathcal{T}} \left| \mathbb{E} \hat{f}_{\tau, n}^{U, \theta_n}(\omega) - \mathbb{E} \hat{f}_{\tau, n}^{U, \theta_0}(\omega) \right| = O(\log n) \|\theta_n - \theta_0\| = o(\sqrt{nb_n})$$

since by assumption $nb_n = o(n^{1-1/(2k+1)})$ for some $k \in \mathbb{N}$. \square

6. Additional plots

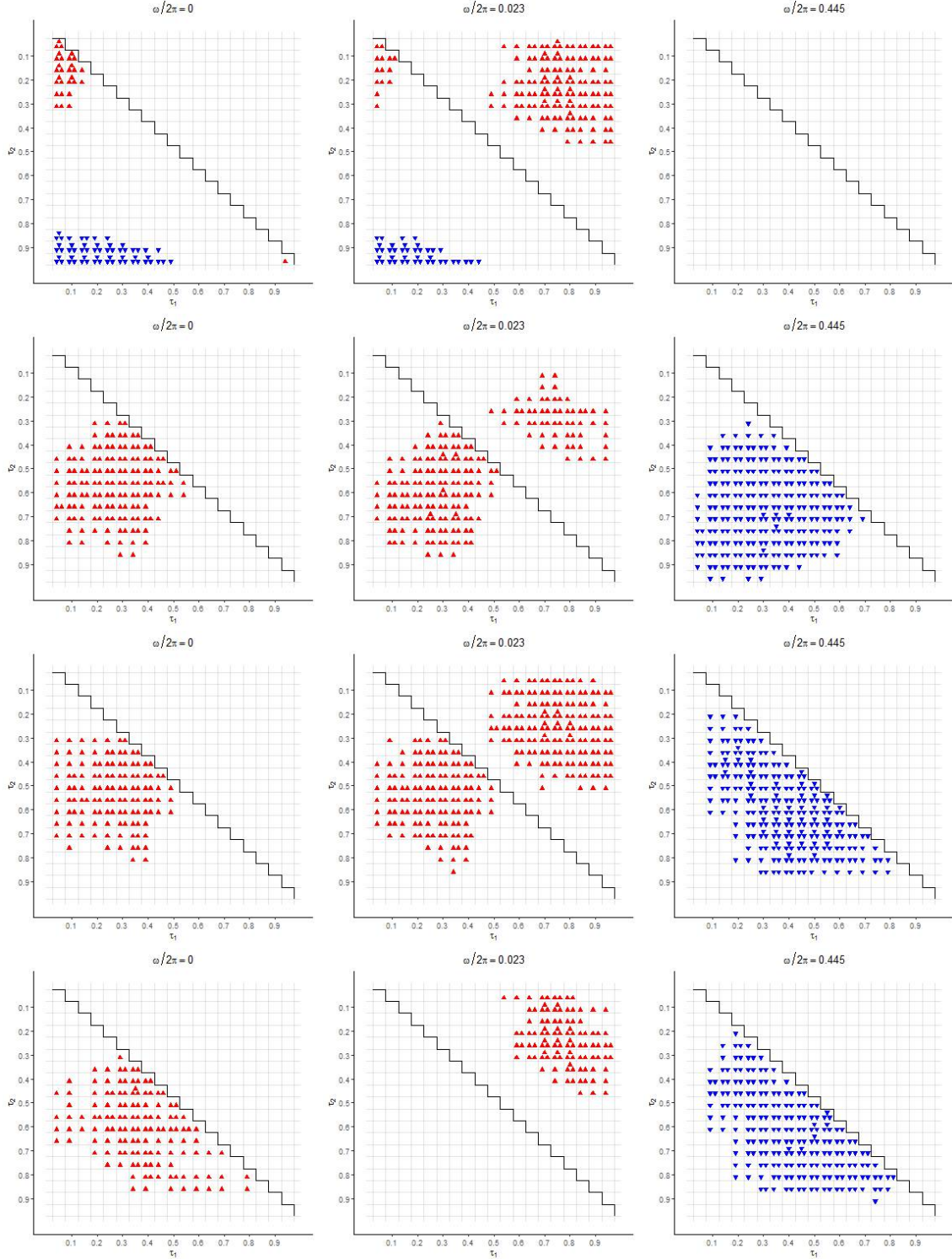


Figure 15: Detailed plots produced by Algorithm 2 at two particular frequencies based on the daily log-returns of the S&P 500 between 1966 and 1970 with AR(3) (top row), ARCH(1) (second row), GARCH(1,1) (third row) and EGRACH(1,1) (bottom row) as candidate model classes.